

*[Handwritten signature]*

# SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**  
\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

## FOR OFFICIAL USE ONLY

*****	
STAFF USE ONLY	
Searcher: <u>PS Dwork</u>	Type of Search
Searcher Phone #: <u>308-4477</u>	NA Sequence (#) _____
Searcher Location: <u>CMI-1B17</u>	AA Sequence (#) _____
Date Searcher Picked Up: <u>6/23/00</u>	Structure (#) _____
Date Completed: <u>6/27/00</u>	Bibliographic _____
Searcher Prep & Review Time: _____	Litigation _____
Clerical Prep Time: <u>4</u>	Fulltext _____
Online Time: _____	Patent Family _____
	Other _____

Vendors and cost where applicable	
STN _____	
Dialog _____	
Questel/Orbit _____	
Dr.Link _____	
Lexis/Nexis _____	
Sequence Systems <u>965302</u>	
WWW/Internet _____	
Other (specify) _____	







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 11:30:45 ; Search time 12463.9 Seconds

(without alignments)  
-522.537 Million cell updates/sec

Title: US-09-180-798-1

Perfect score: 6695

Sequence: 1 tctagatgacgaatcgcgc.....attaatatatttatgtt 6695

Scoring table:

IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_dal:\*  
2: gb\_da2:\*  
3: gb\_com:\*  
4: gb\_ov:\*  
5: gb\_pat:\*  
6: gb\_ph:\*  
7: gb\_pl1:\*  
8: gb\_pl2:\*  
9: gb\_pl1:\*  
10: gb\_pl2:\*  
11: gb\_pl3:\*  
12: gb\_ro:\*  
13: gb\_sts:\*  
14: gb\_sy:\*  
15: gb\_un:\*  
16: gb\_vl:\*  
17: em\_fun:\*  
18: em\_hum1:\*  
19: em\_hum2:\*  
20: em\_in:\*  
21: em\_om:\*  
22: em\_or:\*  
23: em\_ov:\*  
24: em\_pat:\*  
25: em\_ph:\*  
26: em\_pl:\*  
27: em\_ro:\*  
28: em\_sts:\*  
29: em\_sy:\*  
30: em\_un:\*  
31: em\_vl:\*  
32: gb\_htg1:\*  
33: gb\_htg2:\*  
34: gb\_in1:\*  
35: gb\_in2:\*  
36: em\_dal:\*  
37: em\_da2:\*  
38: em\_hum3:\*  
39: em\_hum4:\*  
40: gb\_pl4:\*  
41: gb\_htg3:\*  
42: gb\_htg4:\*  
43: gb\_htg5:\*  
44: gb\_htg6:\*

45: gb\_htg7:\*  
46: em\_htg1:\*  
47: em\_htg2:\*  
48: em\_htg3:\*  
49: em\_hum5:\*  
50: gb\_pl3:\*  
51: gb\_pl5:\*  
52: gb\_htg8:\*  
53: gb\_htg9:\*  
54: gb\_htg10:\*  
55: gb\_htg11:\*  
56: gb\_htg12:\*  
57: gb\_htg13:\*  
58: gb\_htg14:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	6691	99.9	5 A67796	A67796 Sequence 1
2	560.2	8.4	5 DC093048	U93048 Daucus caro
3	560.2	8.4	5 A67797	A67797 Sequence 2
4	387.4	5.8	5 A67815	A67815 Sequence 20
5	383.2	5.7	8 F23M19	AC007454 Arabidops
6	378.6	5.7	8 AC012654	AC012654 Arabidops
7	317.6	4.7	7 ATEF17M5	AL035678 Arabidops
8	313.6	4.7	5 AC006436	AC006436 Arabidops
9	250.4	3.7	5 A67827	A67827 Sequence 32
10	217	3.2	7 AB013395	AB013395 Arabidops
11	208.2	3.1	7 AB025639	AB025639 Arabidops
12	191.6	2.9	8 AC002292	AC002292 GenomIc s
13	191.2	2.9	8 F17123	AF16182 Arabidops
14	178.4	2.7	7 AB005234	AB005234 Arabidops
15	176	2.6	5 AC005170	AC005170 Arabidops
16	167.4	2.5	7 AB012245	AB012245 Arabidops
17	146.8	2.2	8 SBRK1	Y14600 Sorghum bic
18	117.8	1.8	8 ATAC009755	AC009755 Arabidops
19	117.8	1.8	8 ATAC011664	AC011664 Arabidops
20	116	1.7	8 AC002330	AC002330 Arabidops
21	110.8	1.7	8 ZMA010166	ATJ010166 Zea mays
22	110.2	1.6	8 ATEF84	AL034567 Arabidops
23	110.2	1.6	7 ATEF10M6	AL021811 Arabidops
24	109.6	1.6	7 AB009056	AB009056 Arabidops
25	108.2	1.6	5 E12705	E12705 Arabidops
26	108.2	1.6	8 ATU47029	U47029 Arabidops
27	107.2	1.6	5 E12706	E12706 Arabidops
28	107.2	1.6	7 ATHER	D83257 Arabidops
29	107.2	1.6	50 AC004484	AC004484 Arabidops
30	106.2	1.6	7 AB007644	AB007644 Arabidops
31	105.2	1.6	8 ATU55875	U55875 Arabidops
32	105.2	1.6	44 AC018364	AC018364 Arabidops
33	105.2	1.6	50 AC008262	AC008262 Genomic s
34	103	1.5	8 AB019235	AB019235 Arabidops
35	100.8	1.5	8 F20P5	AC002062 Sequence
36	98.4	1.5	50 AC005312	AC005312 Arabidops
37	98.6	1.5	7 RICEPK1	L27821 Oryza sativ
38	95.6	1.4	8 AF024650	AF024650 Arabidops
39	95.6	1.4	7 AT284202	Z84202 A. thaliana
40	95.6	1.4	50 AC006072	AC006072 Arabidops
41	94.2	1.4	8 ATU95973	U95973 Arabidops
42	94	1.4	8 AC007583	AC007583 Arabidops
43	93.6	1.4	7 AB025636	AB025636 Arabidops
44	93.6	1.4	8 AF058826	AF058826 Arabidops
45	92.8	1.4	8 T32N4	AF162444 Arabidops

## ALIGNMENTS



**RESULT 1**  
**LOCUS** A67796 6695 bp DNA PAT 05-MAY-1999  
**DEFINITION** Sequence 1 from Patent WO9743427.  
**ACCESSION** A67796  
**VERSION** A67796.1 GI:4756622  
**KEYWORDS**  
**SOURCE** Carrot.  
**ORGANISM** *Daucus carota*  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Asteridae; euasterids II; Apiales; Apiaceae; *Daucus*.  
**REFERENCE** De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.  
**AUTHORS** PRODUCTION OF APOMITIC SEED  
**TITLE** Patent: WO 9743427-A 20-NOV-1997;  
**JOURNAL** CIBA GEIGY AG (CH)  
**FEATURES** Location/Qualifiers  
 source 1..6695  
 /organism="Daucus carota"  
 /db\_xref="taxon:4039"  
**BASE COUNT** 1845 a 1182 c 1243 g 2421 t 4 others  
**ORIGIN**  
 Query Match 99.9%; Score 6691; DB 5; Length 6695;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 6695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 acctgggacaaagaagcgcctgatgaactgaattaaaccggtttattattacac 720  
 DB 661 acctgggacaaagaagcgcctgatgaactgaattaaaccggtttattattacac 720  
 QY 721 ttgatgcttcataagactaatactcttggttaattcattgataataacccttcct 780  
 DB 721 ttgatgcttcataagactaatactcttggttaattcattgataataacccttcct 780  
 QY 781 ctgtgatggcagtgagccctcggttaattagcgaactaatactatctgatactgta 840  
 DB 781 ctgtgatggcagtgagccctcggttaattagcgaactaatactatctgatactgta 840  
 QY 841 tgcctggcagaatgatactaatcacaatttgctgtcacaataaataaacc 900  
 DB 841 tgcctggcagaatgatactaatcacaatttgctgtcacaataaataaacc 900  
 QY 901 tctattatgtcatcttttttgtaacaagtaactaattgacgtatgtaacagcg 960  
 DB 901 tctattatgtcatcttttttgtaacaagtaactaattgacgtatgtaacagcg 960  
 QY 961 atgcattacacaacttaagaaactagcttgcagaatcccaaatgtcctgcagagctgg 1020  
 DB 961 atgcattacacaacttaagaaactagcttgcagaatcccaaatgtcctgcagagctgg 1020  
 QY 1021 atccaacccttgtaacccttgacatggtttcaatgtgacatgatacaatgaacaagtg 1080  
 DB 1021 atccaacccttgtaacccttgacatggtttcaatgtgacatgatacaatgaacaagtg 1080  
 QY 1081 ttaataagatgtatgacactcccttataatttttagcaagttacgaataattac 1140  
 DB 1081 ttaataagatgtatgacactcccttataatttttagcaagttacgaataattac 1140  
 QY 1141 aattggcagaatgctccttaataattttcttaattttcttagcgaagcgagatccta 1200  
 DB 1141 aattggcagaatgctccttaataattttcttaattttcttagcgaagcgagatccta 1200  
 QY 1201 tcttaagatcctactgaatttaagacataatacttttttaaaaaatcattagagt 1260  
 DB 1201 tcttaagatcctactgaatttaagacataatacttttttaaaaaatcattagagt 1260  
 QY 1261 gtttttccgacagcgacataatacttttttgtaattcaagaacaccttcctcg 1320  
 DB 1261 gtttttccgacagcgacataatacttttttgtaattcaagaacaccttcctcg 1320  
 QY 1321 acctaaataataagatataactccttgaaactaatttttaattttcttcttcta 1380  
 DB 1321 acctaaataataagatataactccttgaaactaatttttaattttcttcttcta 1380  
 QY 1381 tgtctcttcagagaagttcttaatgcttttgtaaaagtacattctatgaataattt 1440  
 DB 1381 tgtctcttcagagaagttcttaatgcttttgtaaaagtacattctatgaataattt 1440  
 QY 1441 tggcaactcatalaataattatataatccatgtagtlaabaagttaaaaaaccttccta 1500  
 DB 1441 tggcaactcatalaataattatataatccatgtagtlaabaagttaaaaaaccttccta 1500  
 QY 1501 ttaattccaagaatagaggtcattttatagtttggaacccacagtggttttgaanaatgt 1560  
 DB 1501 ttaattccaagaatagaggtcattttatagtttggaacccacagtggttttgaanaatgt 1560  
 QY 1561 cagaatcttgtagtlaatttacttaccacactttatagcgatcagtgatcttg 1620  
 DB 1561 cagaatcttgtagtlaatttacttaccacactttatagcgatcagtgatcttg 1620  
 QY 1621 ggaatgcagaattactgtgcaattggttcctcttgccagttgaaanaattacaact 1680  
 DB 1621 ggaatgcagaattactgtgcaattggttcctcttgccagttgaaanaattacaact 1680  
 QY 1681 tgaagaacatataacttgtaagtgcttagtttttaagcgcaaatgctttaaatct 1740  
 DB 1681 tgaagaacatataacttgtaagtgcttagtttttaagcgcaaatgctttaaatct 1740



QY	1741	gttaaaagtgtgaaaaagtgacttcttcgttcgaagcgtgttggaataactatga	1800
Db	1741	gttAAAAgtgtGAAAAgttgAACTTCTTCAGCTTCACCAAGTTTGTGGAAATATCTAAGA	1800
QY	1801	agcaactaaaggctcgggcaactttcttgtaattcaatacttgtaattgta	1860
Db	1801	AGCACTTAAAGGCTGGGCAATTTTTTGTATATATTCAAATATGTGTAATGTACTA	1860
QY	1861	cttaatatgtaaacttgatttaactccctcaatgatggtctcagttccaaatgtgcctcaat	1920
Db	1861	CTTAATATGATTAAGCTGATTTAACTCCATGATNGGTCTCAGTCCAAATGAGCCOTATY	1920
QY	1921	agtcacactaataaatttgnrgggtgtggacaataataactcttcttctaaggtccagaa	1980
Db	1921	AGTCACACTAATAAATTTGGNGGGTGGACAAATATTAACCTCTTCTTAAAGTCCAGAA	1980
QY	1981	gagcacatacaacctgtctcagcgaatacgtcaagcagcaggtgcagtcagcagcattcca	2040
Db	1981	GAGCACATTATCAACCTGTCTACGCGATACGTACAGTGGGTCAAGTCAAGGCTATCCA	2040
QY	2041	gtttgggaaggtttaatgagcacattattacgtgtctttaaagctctgagagtgtta	2100
Db	2041	GTTTGGGGAGGTTTATGATGAGCACTATTATTAACCTGTCTTAAAGCTGAGGATGTTA	2100
QY	2101	ttaagcgtcgtcatcatcctcagaggtttaattagcacttcagtgtctatataatgaagtaca	2160
Db	2101	TTAAAGCTCAGATCATTCACAGAGTTTAAATTAGACCTTCAGTTGATTAATGAATGGTACA	2160
QY	2161	tgaagaatagacatacttaaggttcceatgcgtgtttcaacatgctcctaatectcgtta	2220
Db	2161	TGAAGAATACATATCTTTAATGTTCTCATGCTGTTTCAACATGCTCTAATATTCGTGTA	2220
QY	2221	tctttgcatctttaaaaaatgycacgttgattaaaatgtagaagaagtagcttccaaatacca	2280
Db	2221	TCTTTGCATCTTAAAAATGGCAGCTGATTTAAATGGAAGAAAGTATTCATTCAAATACA	2280
QY	2281	tttaagtataccagagaaataatcaataattttttaatcaataagttggccctagagttc	2340
Db	2281	TTTATATATATACAGAAATATCATATTTTAAATCATATAGTTGGCCCTAGAGTTT	2340
QY	2341	tctcagatgttggtcattattatcttccacactttagaacgtgtgttcagatgaaatac	2400
Db	2341	TCTCAGATTTGGTCTATTTATATTTTCCACATTTAGAACTGTGTGTCAGTGAAGAAATC	2400
QY	2401	ttgagacttccacagaagaatctttaagtaaagtaattctttaagctgataagtaagtg	2460
Db	2401	TTGGACTTCCACAGAAAGATCTTATAGTAAAGATTCCTTAAATCTGATGATGAAGAAAGT	2460
QY	2461	tcaatgtgtggtcgtgtcccgaaatttaaatcaatcccatgcaacatggttgtatctga	2520
Db	2461	TCAATGAGTGTGGCCGTGCCAGAAATTTAAATCAATCCCATGTCACATGTTTGTGATCTGA	2520
QY	2521	ctactcaactgtttaatcgaaagaatcaactatttfggaatgaatgaagcctttttttgtcttc	2580
Db	2521	CTACTCACTGTATATCGAAAGATTAATTTGGAAATTAAGCTTTTGTGTTT	2580
QY	2581	catgcttagcgtttaaaaggtctacgtctcgaactatggttttaacaatgtaataagttgt	2640
Db	2581	CATGCTTAGCGTTATAAAGGCTACGTCAGTCAGCAATGCTTTAAACATGTATAGTTTGT	2640
QY	2641	actgacaagaatttaaaagttctctctgtttaagaaatgaagaataataataaagcctt	2700
Db	2641	ACTGACAGAAGTTTAAAGTTTCTCTGTGTTACGAATTTAAGAAATATATAATATAAAGCTT	2700
QY	2701	aacttcctcgttggaaggtgtcttcaacttttatataatatagatattcaagacctc	2760
Db	2701	AACCTTCTCTGTGGAAGGTTCTTACCTTTTATATATATATATAGAACCTCACAGACTCT	2760
QY	2761	gctggcaattatatacttaacgaacttcagatatacagaagaactgtatataagttcaagt	2820
Db	2761	GCTGGCAATTATATCTTAGACACTTACGAGTATACAGAACTTGATATATTAAGTTCAAGTG	2820
QY	2821	agtggtcgtgtagaacaccttaagcaagaacttaacatcagtgagttcaaacctttaact	2880

Db	2821	AGTGGCTGTAAGTGAACACCTTTAGACAAACTTATATCATGAGGTTTAAACCTTTTAACT	2880
QY	2881	ttccttttaagatttttttccaagtttacygaanaatttaccctcgtacgtgtttctttcc	2940
Db	2881	TTCTTTTAACTTTTTCACAGTTATAGGAAATTTGACTCTATGATGCTGGTTCTTTC	2940
QY	2941	ataaacttcacataaagtcggtttcttctgaagtttcabyltaagctgtttgcagatgt	3000
Db	2941	ATAAACCTTCCATATTAAGTCCGTTCTTGTACGTTTTCATGATGATGCTGTGACGATGAT	3000
QY	3001	attagcggtttcttccaataatctcaatgtgtctcaactttgatagagcctgtactattat	3060
Db	3001	ATTAGCGGTTCTTTCAATTAATCATTAATGCTGTCACTTTGAGAGCGCTGTACTATTAT	3060
QY	3061	tgcaacttgcacttaaaccttgatccccaatcgtatcattctgaattgtcctagctaccacga	3120
Db	3061	TGCACCTTGCACTTAACCTTGATCCCTCATGTCATCTTGATTTGCTATGCTACTACCGCA	3120
QY	3121	gtcgacaatggtttatcaatgctcttttgaagtgaaacaatgtagctttcaactctgtccttga	3180
Db	3121	GTTGACATGCTTTATCATGTCCTTTTGAGGTAAACATGATGATGCTTACCTCTGTCTGGA	3180
QY	3181	tataggttttaagccttgcacactcccaactagccttccttggtttattcaacagtccaaca	3240
Db	3181	TATAGGTTTAAAGCTTTCACCTCCACTCCACTGACCTTTTCGTTTATTCACAGTTACACACA	3240
QY	3241	ccctacagcaactgttcaacccctctagtccttttgcgcggaatagtaagaagtttcttcgca	3300
Db	3241	CCCTACTGACACTGTTCACTCCCTCTGTCTTTTGTCCGCAATATGATGAAGATTTCTTTGCCA	3300
QY	3301	taatagtgagatgacatcatttaagaataagtgaaatcaaatatcgtttatgttgtttctac	3360
Db	3301	TAAATAGTGAGATGATCATTTTAAGAAATAGTAATCAAAATATATGCTTATGTGTTGTGAC	3360
QY	3361	tttgggaattaaatgagttcgtctgacaattgtgtcgtttatcgtgtgtccaagcctttgcca	3420
Db	3361	TTTGGAATTAATATAGTTAGTGTGACAACTTGTCTGTATTATCCTTGTCAAGCCTTTGCCAA	3420
QY	3421	ggaagcgcaatagaagaagtgggcaccacaagcgcccttaactcttgaaggggggcgagac	3480
Db	3421	GGAAAGCGAATTAGTAAGATGAGGACATCCAAAGCGCTTTATCTTGAAGGGGGGGGGGAC	3480
QY	3481	gttgtagattcctgggtgtcctattagaggaacatlatcatatatactgattattataga	3540
Db	3481	GTTGTGGAATCTGGGTGTCTATTAGAGGAATATATCTATATATATGATTTTATATAGA	3540
QY	3541	atataaatcaactacatataatlttctttgtaatgttatatagaatcccaactcgtaaac	3600
Db	3541	ATAATAATCAACTCTATATTTTCTTTGTATGTTATATGAAATGCCACTGTATAAC	3600
QY	3601	tttggcaaataccattgnaaatttgaacctaatttaattagttgttgaggttttaattca	3660
Db	3601	TTTGACAAATACCATTTGAAATTTTGAACCTAATTAATTATGATGTGTACAGTTTAAATTCA	3660
QY	3661	aactcatttaatttacttttaaaaaaataattcatabatgtsatcgttacaagttataatat	3720
Db	3661	AACCATTTTAACTTTTAAATAAATAATTTCTATATGAAATGTAACAGTATAAATATAT	3720
QY	3721	taaatatagatgatatgtgtcctatataatagctgnaatgtctaaatagaactccaagaagcgt	3780
Db	3721	TAAATATGATGTATGTGTGCTATATATATAGCTGAATGTCATTAATACATCCAAAGCGCTG	3780
QY	3781	ccttaactagccttaagggttcgaagccaattacagatgagttactacttcaaaatatagggttc	3840
Db	3781	CTCTTACTGCTTAGCGCTCCAGGAGGTCACTGATGCTTAACCTTATACAAATATGGGGTTC	3840
QY	3841	gtatagaatgttttgaggaatcccatcaacatcagatctcgttttgcgcacccctcttcaat	3900
Db	3841	GTAATGACATTTGTTGGGATATCCCATACAGAGATTCCTGTTTGTCTGACCCCTGTCTCAAT	3900
QY	3901	tgatttccaatgatatgattactacgttttataaataatcttatttgcataaatttaac	3960



Db 3901 TGATTTTCATGTAGTAGTATTACTAGTTTATATAATATCTTATTGCAATAAATTTAAC 3960  
QY 3961 tggagattcaaatgacagggagctttacagcaataacataagtgaccattctctgtg 4020  
Db 3961 TGGAGTTTAAACAATGACAGGGAGCTTTACACCAATTAACATAGGAGCAATTCCTAGTG 4020  
QY 4021 atcttgggaatctgcaaatcttgtagcttggaacctataatgaatagctctctgagc 4080  
Db 4021 ATCTTGGGAATCTGCAAAATTTGGTGAAGCTTGAGACCTATACATTAATAGCTTCTCTGAGC 4080  
QY 4081 ctataccggaacacatlaagaaagcttacaaggttaagaattctgtatgactacaactct 4140  
Db 4081 CTAATACCGGACACATTAGAAAGCTTACAAAGCTTAAGATCTTGATGACTACAAATCTT 4140  
QY 4141 caatagttttaactaatgcaatttgatctcttctaagtgtgattatatacacaaa 4200  
Db 4141 CACTAGTTTAACTTAAATGCAATTTGAATTAATCTTCAATGATGATGTTATACACAAA 4200  
QY 4201 ttactggataggcgtctcaacaacaacactgcctctgtgccaattccaatgtcaactgact 4260  
Db 4201 TTACTGGATAGGCGTCTCAACAACAACAGCTCTGTGCTCCAAATTCCAATGTCACTGACT 4260  
QY 4261 aatattacaactctcaagctccgtlaaglatccgaccttccagatagtttgtgtg 4320  
Db 4261 AATATTACAACCTCTCAAGTCCGTGAAGTATCCGACCTTCCAGATAGTTTGTGTG 4320  
QY 4321 tggatgtttcaatttaactaataatgtaactgaagattacaagaatgacctat 4380  
Db 4321 TGGATGTTTCAATTTTAAATTAATTAATGATCAAGGATTTTACAAACAATGGGCTAT 4380  
QY 4381 caggacacagtaacagataatgagctcattctctgttacaacctacaaggttaagtctta 4440  
Db 4381 CAGGACACAGTACGGATATAGGCTCATTTCTTGTTTACACCTATCGATTAATGCTA 4440  
QY 4441 gtaatactttaataattatggtcttactctactctacgtcgaaagctatgataattcttc 4500  
Db 4441 GTAAATATCTTAAATTAATTAATGTTCTTACTTACTACGCAAAAGCTAATGATTAATTTT 4500  
QY 4501 tctctctcaataataactctgagtttggaataattgaattatgtagactg 4560  
Db 4501 TCTCCTTCATTAATTAATCACTTTCGCAAGTTTGGCAATTAATTAATTAATGAGACTG 4560  
QY 4561 taacttggagagccctgagctcagatctcccaattctcccaacctccagcttcaaccac 4620  
Db 4561 TAACTGGGAGGCGCTGCGCTGATCTCCCAATTTCTCCACCAACCTCGTCAATCCAC 4620  
QY 4621 catcaacagtaacagctccaggtgattagttttatataatcaccgtaattactta 4680  
Db 4621 CATCAACAGTACAGCTCCAGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4680  
QY 4681 tgactgttaaaaaattggttaattccacagttgcaataaagtaatttctctctctc 4740  
Db 4681 TGAATGTAAAAATTTGGTGAATTTCCACCAAGTGGCAATTAAGTATTTCTTCTTCTC 4740  
QY 4741 tttctattatlaagaagacaatggtcccaactgagactatctgagggagagtagctg 4800  
Db 4741 TTCTATTATTAATGAAGCAAAATGTCCTGACCTGAGCTATGCTGAGGAGTAGTCTG 4800  
QY 4801 tgggtgctgttactgattgtctgcacccgcaatgagcaattgcaatggtggagagaagaa 4860  
Db 4801 TGGTGTGCTTTACTGTGTGCTGCACTGCAATGGCATTTGCAATGGTGGCGAGAAAGAA 4860  
QY 4861 accgagagaacattcttggatggtccaggttagtccgttaataatgatactattagaagc 4920  
Db 4861 ACCGGAGAAACATTTCTTTGATGTGCCAGGTAGTCCGTAAATTAAGATATCTATTGAAGC 4920  
QY 4921 gcttactgtctgtaggacttgttttcaactgltcaatgattgaattcaactgagagaagccc 4980  
Db 4921 GCTTACTGTCTGTGAGCTTTGTTTACAGTATAGTAACTTCACTGAGTGAAGAGGAGCC 4980  
QY 4981 agaaagagcccttgttcaactgaaagagtttctctcgagagaattgcaagctcgcaagga 5040  
Db 4981 AGAAAGTACCTTGTGCAACTGAAGAGGTTTCTCTGCGAGAAATGCAAGTCCGAAACGA 5040

QY 5041 tacttttagtaccatcccttggagaagagtgatttggtaagtgatagaagagcccttgc 5100  
Db 5041 TACTTTTAACTACATCCCTTGGAAAGGTGGATTTGGTATGATTAAGGAGCCCTTGC 5100  
QY 5101 tgatgctcaacttgaagcaattaaaggcttaagaagaagaacacaggttgcagact 5160  
Db 5101 TGATGGCTTACTGTATGACAGTTTAAAGGCTTAAGAAACGAACACAGGTGGCGAGCT 5160  
QY 5161 gcaattcaacaagaagtggaatgattagcaatggtgtgcatcgaaatctctgtgct 5220  
Db 5161 GCAGTTTCAACAGAGTGAAGTGAATGATTAGCATGCTGTGCATGGAATCTTCTGCGCT 5220  
QY 5221 acgtgtttctgcatgaacacctgaagcagcgtctcttataatcaatacaggttaatg 5280  
Db 5221 ACGTGTTTCTGCATGACACCTTACCGAGCGGCTCTTGTATATCCATACATGCGCTAAATGG 5280  
QY 5281 aagtggtgcatgttttaagaggtatctcagttacaattacaacttgcagaaatt 5340  
Db 5281 AAGTGTGCTCATGTATTAGAGGTATCTCAATTACATTAACATTCAGAGAT 5340  
QY 5341 tgttgaattaaaaatgaatatlaactccctacatagttgaagtgtlaataattctgag 5400  
Db 5341 TGTGTATTAAAAATGAATATTAATCTCCCTACACTATGTAAAGGTATTAATTTCTGAG 5400  
QY 5401 cagatctatctccattgcaagatacagatattattgttttctgttaattgatacgg 5460  
Db 5401 CAGATTTATTTCCATTCGACAGATACCAATTAATTAATGTTTCTGTATTAATACG 5460  
QY 5461 gtataattcttcttctgtlaatttggatatacgaagatttcgagttctaataatgataca 5520  
Db 5461 GTATATTTCTTCTTGTGTTATGTTATGCAAGATTTGAGTCAATTAATTAATTAATTAAT 5520  
QY 5521 actgagctcattgattatctcgaattgaattctgtctcaatggtgccaatatatga 5580  
Db 5521 ACTGAGTGTATGTTATTTATCTGCAATTAATTTCTGCTTCAAGCCAAATTAATTAATGA 5580  
QY 5581 ttcaatttggaaatctataataataactggtgtaagtcagctgttgaattcttcaatta 5640  
Db 5581 TTCAATTTGGAATCATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5640  
QY 5641 attagttctcaataacagaactctgcagtagagcttaccgacatactctaacccttc 5700  
Db 5641 ATTAGTCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5700  
QY 5701 ttatggccctgatatataatcgtcccaacttacttattgaatttgcctctcgaattc 5760  
Db 5701 TTATGGCCCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5760  
QY 5761 ttgatctgtaacttggatgcttcttcaatcaaatgtaagcgtcagacatcagaagcc 5820  
Db 5761 TTGATCTGATCATGTGATGCTTGTGTTTCAATCAATTAAGCGTCAAGCATCAAGAAC 5820  
QY 5821 tccccctgattggccaactaggaagagattgacactagatcttctaggggactctaa 5880  
Db 5821 TCCCCCTGATTGGCCAACTAGGAGAGATTTGCATAGGATCTTCTAGGGCCCTATCTAA 5880  
QY 5881 attgcatgaccatttggatcccaaatattacatctgagatgttaaaagttgaaatatatt 5940  
Db 5881 ATTGCAATGACCATGTGATCCCAAGATTAATTCATGCGCATGTAAGTGAATTAATTT 5940  
QY 5941 attgagcagaagaatttgaagcgttctgtaagttatttggtagttagcccaatgagatta 6000  
Db 5941 ATTGAGCAGAAATTTGAGGCTGTGTAGGTATTTTGGGTTAGCTAGGCTCATAGATTA 6000  
QY 6001 caagatacccatgtaacagctcgttaagggtaaccatttgggacataagctcccagta 6060  
Db 6001 CAAGGATACCATGTTACAGCTGTGTAAGGGGTACCATTTGGGACATAGCTCCGAGTA 6060  
QY 6061 cctctgactgtaagattctcagagaagccgattcttggttatggataatgctcct 6120  
Db 6061 CCTCTGACTGGAAGTCAATCAGAGAACGATGCTTTGGTTATGGGATTAATGCTCCT 6120



QY	6121	aaagacattactacggacaagaagcctttgacactgtctgcgccttcggacacataat	6180
Db	6121	AGAGCTCATTACTGGACAGAGGCGCTTTTGAATCTTCTGCGCTTTCGACACATATATG	6180
QY	6181	tatgttgtagttagttagttagtgcgccggagtcttcctttagttaaattattccacatta	6240
Db	6181	TATGTTGTTGGATTGGGATGTGTGCCGGGTGTCCTTGTGATTAATTTACATATTA	6240
QY	6241	gtctctacacacttggttggtagcccttctttttattcttcgcgcgtattatgtctagt	6300
Db	6241	GTCGTTACRACCTTGTGTGGGCCCTTTGTATTTCCTCCGATTCGATTTGATTTTAA	6300
QY	6301	catgttatgcataatgtagcctgttcgaatgctttagtggttaaaagccctttgaaagag	6360
Db	6301	CATGTTATGCAATTCACCTGCTTTCGAATGCTTTTGTAGTTAAAGCCTTTGAAAGAG	6360
QY	6361	aaaaagcttgagagatgctgtcgcatactcgaacccgcgaacaaattcatgtgacaaagat	6420
Db	6361	AAAAAGCTGAGACTCTGTGTCATCTCTGACCTCGCAAGAACATTTCAATTGACAGAA	6420
QY	6421	gagcagcttattaaagtagcattactctgtacccagaaggtctgccaatgagcggctaa	6480
Db	6421	GAGAGGCTTATTCATACATCACTTACTGTGTACCCAGGGGTGCGCAAGGAGACGGCT	6480
QY	6481	atgtcacagatagtcgcgaatgctttaaaggtatgacgttccttgagaaagatgagcag	6540
Db	6481	ATGTCAAGATTAATCCGAATGCTTGAAGGTATGAGCCTTCAGAAAGTGGAGACAGT	6540
QY	6541	caaaaagtgaagtcataccatcaagaacgtagaattagctccacatgcgaactctgaag	6600
Db	6541	CAAAAAGTTGAATCTATCCATCAAGACGTAGATAGCTCCACATCGAATCTTGATG	6600
QY	6601	atcttagaccgcgaacataactgtcagcttttgaattattgtgtccaaagataaagaca	6660
Db	6601	ATCTTAGACTCGACACATATACCTTGATGCTTTTGATATATGTGTCACAAATTAAC	6660
QY	6661	tataaatgtatgaataataatttttttagtt	6695
Db	6661	TATATAATGTAAATGAATATATATTTTATATGTT	6695
RESULT	2		
DCU93048			
LOCUS	DCU93048	1755 bp	mRNA
DEFINITION	Daucus carota somatic embryogenesis receptor-like kinase mRNA,		28-JUN-1997
	complete cds.		
ACCESSION	U93048.1	GI:2224910	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
	Daucus carota		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
	eudicots; Astceidae; euasterids II; Apiales; Apiaceae; Daucus.		
REFERENCE			
AUTHORS	1 (Pases 1 to 1755)		
TITLE	Schmidt, E.D., Guzzo, F., Toonen, M.A. and de Vries, S.C.		
	A leucine-rich repeat containing receptor-like kinase marks somatic		
	plant cells competent to form embryos		
	Development 124 (10), 2049-2062 (1997)		
JOURNAL	97313247		
MEDLINE			
REFERENCE	2 (Pases 1 to 1755)		
AUTHORS	Schmidt, E.D., Guzzo, F., Toonen, M.A.J. and de Vries, S.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-MAR-1997) Molecular Biology, Agricultural University		
	of Wageningen, Drievleaan 3, Wageningen 6703 HA, The Netherlands		
FEATURES			
source			
	1. 1755		
	/organism="Daucus carota"		
	/db_xref="taxon:4039"		
	34. 1695		
	/note="SRK"		
	/codon_start=1		
	/product="somatic embryogenesis receptor-like kinase"		
CDS			

Query Match	Best Local Similarity	8.4%	Score 560.2	DB 8	Length 1755
Matches 735	Conservative	0	Mismatches 18	Indels 144	Gaps 2
QY 5800	agaagcgcacgacccatcagaaaccccccctgattggcccaactiagagagagattgcaatag	5859	DB 984	AGAGCGCAACCATCGAAGCTCCCTTGATTGGCCAATAGAGAGGATTCACATAAG	1043
QY 5860	atctctcagagagccatactcctaattcagaccatttgatccccaagttatcatatgca	5919	DB 1044	ATCTGCTAGGGGCTTTCTTATTTCATGACACCATTTGATCCCAAGATTATCCATGTGTA	1103
QY 5920	tgtaaaagctgcacaatataatttgcagcgaatattgcagctgtgtgtaagtgatttgg	5979	DB 1104	TGTAAAGCTGCAAAATATATATATGAGAGGAATTTGAGGCTGTGTAGCTATTGTTGG	1163
QY 5980	gttagcaggagctcattgattacagaagatcccatgttagactgctgtgaagggttacat	6039	DB 1164	GTTTAGCTAGGCTCATGAGATTACAGAGATACCATGTTTCAACTGCTGTGAAGGTTACTTT	1223
QY 6040	tgaggacatagtccccagtlacccctgcactgcagtcgaagtcacatcagaagaacgcga	6099	QY 6100	tggtatcagagataatgtcctcctagagcctattaccgcagacagagagccttattacttcg	6159
DB 1224	GGGCTACATATGCTCCCGAGATACCTCTGACACGGAAGTACATCAGAGAGACGATGTCTT	1283	DB 1284	TGCTTATGGGATTATCTCTTAGAGCTCATTAAGTGAAGAGAGACCTTTATCTTGTCTGG	1343
QY 6160	cccttgacagcatgatgatgttaactgtgttgagattgggtatgctgccgggtgttcctt	6219	DB 1344	CCCTGACCAAGATGAGATGTTATGTTGTGATGTTG-----	1380
QY 6220	ggttaattattcacatatagtagtgcttaactattgtgttgccccctgtgtttattcc	6279	DB 1380	-----	1380
QY 6280	tgacctgatttgatttccttgatcattgataattgcataattgcactgctgtgcaatgtcttag	6339	DB 1380	-----G	1380
QY 6340	gttaaaagccttltgaaagagaaaagtgtgagatgctggtgcagtcctcagcctgcagaa	6399	DB 1381	GTTTAAAGCTTTTGAAGAAGAAAAGTTGAGAGATGCGTGTGATCCTGATCTGAGAAC	1440
QY 6400	aattacattgcacacagaagtgtgaagcagcttatacaatgaacattacatctgttaccaga	6459	DB 1441	AATTATCTTATACACAAAGATTGAGCAGCTTATTCAAAGTAGCATTAACCTGTATACCGAGGT	1500
QY 6460	tcgcgaattgagagcctcgaagatgtcagaagtgatccgagtccttgaagtgatgacct	6519	DB 1501	TCGCGCATGAGCGCGCTTAAGATGTGACAGGTAGTCCGAATGCTTGAAGGTGATGCGCTT	1560
QY 6520	gcgaataaagtgaggagcagtggtgcataaaagtgtgaatgcatccatacgaagcgtgaatagct	6579	DB 1561	GCAAPAAAGTGGGACAGTGGCAAAAAGTAGAAGTCAATCCATCAAGACGTATGAATTAAGCT	1620
QY 6580	ccaatctgaaccttctgaatgagctcctagactcagacagaagataactctgcatgtctttgaatca	6639			



OY	6640	tctgtccagaataaacagcatatataaatg-t-aatgaatatatttttttgattt 6695
Db	1681	TCTGTGTCACAGATTAACAGCATATTAAATGTGAATAAATAATTTTATAGGTT 1737
RESULT	3	
A67797		
LOCUS	A67797	1815 bp DNA PAT 05-MAY-1999
DEFINITION	Sequence 2 from Patent WO9743427.	
ACCESSION	A67797	
VERSION	A67797.1	GI:4756623
KEYWORDS		
SOURCE	.	
ORGANISM	carrot.	
	Daucus carota	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	euphyllophyes; Spermatophyta; Magnoliophyta; eudicotyledons; core	
	eudicots; Asterales; euasterids II; Apiales; Apiaceae; Daucus.	
REFERENCE	1 (bases 1 to 1815)	
AUTHORS	De V.S., Schmidt E.D., Van.H.G. and Hecht,V.F.	
TITLE	PRODUCTION OF APOMITIC SPED	
JOURNAL	Patent: WO 9743427-A 20-NOV-1997;	
	CIBA GEIGY AG (CH)	
FEATURES	Location/Qualifiers	
source	1..1815	
	/organism="Daucus carota"	
	/db_xref="taxon:4039"	
	94..1755	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAB42241.1"	
	/db_xref="gi:4756624"	
	/translation="NNRNSILNLYMOFTDAYLDKRYGLMTLELXSNNGIPSPSDLG	
	NLTNLVSVDIYNSSSGPIPTDLKLRLRFLRLNNNSISGPIMPSTNTITLOLDL	
	NNNRISGPVPNGSPSLPTPLISFANNLNLGGPVGRCGSQPSPPSPPTSPVOE	
	PQONGPTGAIAGVAAGAALLFAAPAMAFPMWRRRPRREHFDVPAEDPEVHGQL	
	RSLSELRVALDTSTTLIRGRGFGRVTKGRLAVAKRLKEETRGEGLOFOTEL	
	LISLSAIVHNLLRLRNGFCMPTERLLVPPFMANGSVASCLRQDPSEPPLDMPTKR	
	AVGSRAGSLYHDHCDPKIHRDYKANIIIDDEFRAVVDGLANDYKDHTTTR	
	VETSLGYLAPEYLSTGSKSEKTDFVEGIMLDELIGORAFDIARLANDDDVMLDMW	
	KSLTEKKLEMLVDPLENNYIDTFEEVOLIAVLCTGCSPMERPMSEVVMLEGDC	
	LAEKDENQKVHYDVIELAPRHRTSMILDSIDLNAHELSPSR"	
BASE COUNT	531 a 354 c 415 g 515 t	
ORIGIN		
Query Match	8.4%; Score 560.2; DB 5; Length 1815;	
Best Local Similarity	81.9%; Pred. No. 1,5e-86;	
Matches 735:	Conservative 0; Mismatches 18; Indels 144; Gaps 2	
OY	5800	agagcgtcacgcacacagaccctcccccgattggccaactaggagagagattgcaatag 5859
Ddb	1044	AGAGCGTCAGCCATCAAGAACCTCCCTGATTGGGCCCACTAGSAAGAGATTGCACTAGG 1103
OY	5860	atcttcaggaggacctctaactatgatgaacattgtaaccatcccaagattatccatcgca 5919
Ddb	1104	ATCTGCTAGGGGGGCTTTCTATTATTGCAIAGCCATTGTGATCCCAAGATTATCCATGTA 1163
OY	5920	tgtlaaaagctgcacaatatattatttgacagaaagtatttagagctgttgttagtgatttg 5979
Ddb	1164	TGTAAAAGCTCAAAATATATTATTATTTGGCGCAAGAAATTGAGGCTGTTAGAGTGATTTGG 1223
OY	5980	gttagctaggtcatatggaattacaagaataccatglttagagactgctgtgaagggttacat 6039
Ddb	1224	GTTACTACTAGGCTCATGTGATTACAAGGATACCATTGTTACAACTGGTGTGAAGGGATTACTT 1283
OY	6040	tgggacacatagctcccagafacctcttcagacttgaaaagtcatatagagaagacagatgctt 6099
Ddb	1284	GGGCTACATTAAGTCTCCCAAGTACTCTTCGACTGSAAGTATATAGAGAAGACGATGTCCTT 1343
OY	6100	tggttatgggaataatgctcctcagaatcattactatctgagacagagagcttttgattcttgctg 6159

	Dy	6160	ccttggaacagatgatgatatgttatcttcttgattggtagtatagtgtcccgagtctccttt	6219
	Dy	1344	TGGTTATGGAATTATAGCTTCTTAGACCTCATTAAGACAGACACTTTTGATCTGTGC	1403
	Dy	6160	ccttggaacagatgatgatatgttatcttcttgattggtagtatagtgtcccgagtctccttt	6219
	Dy	1404	CCTTGGAAACGATGATGATGATTTATCTTTGGATGG-----	1440
	OY	6220	ggttaattaattcacataatagtgcttactactctgttgttgcccccttgttttatttc	6279
	Dy	1440	-----	1440
	OY	6280	tgcctgtatttgattccttgatcatgttatgcataatcacccgtcttgcaatgctcttag	6339
	Dy	1440	-----g	1440
	OY	6340	gttaaaggccttttgaagaagaaaaagttgagatgctgcgtcgcatacctgcagacaac	6399
	Dy	1441	GTTAAAAAGCCTTTTGAAAAGAAAAAGTTGGAGATGCTGTGCAATCCCGAACCCTGGAAAC	1500
	OY	6400	aattacattgcacacagaagttgagaagcttattcaagtagcatactctgtaccagggt	6459
	Dy	1501	AATTACATTGACACAGAAGTTGAGACAGCTTATTCAGTAGACATTAACCTGTGACCAAGGT	1560
	OY	6460	tcgccaatgagagggcgctaagatgtfccaagaggtatgcgaatgcttgaaggtgtagcgctt	6519
	Dy	1561	TCGCCAATGAGAGGGCGCTTAGATGTCAAGAGGTATGCCAATGCTTGAAGGTGATGGCCTT	1620
	OY	6520	gcagaaagtgygagcagatgycacaaaagttgaagtcatalccataagacgtataagttgct	6579
	Dy	1621	GCAGAAAAGTGGGACGAGTGGCGAAAAGTAGAAGTCAATCCATAAGACGTGAATTAGCT	1680
	OY	6580	ccaatctgcgaactctctgaatggtatcccaagctgcagagtaactgtcagctttgaatia	6639
	Dy	1681	CCAAATCGAACTTCTAAAGGATCTGAGCTCGACAGATPACTTGCAAGCTTTGAATTAT	1740
	OY	6640	tcgtgtccaagatcaaacagcataataaagtt-aatgaaattaatattttttatggt	6695
	Dy	1741	TCGTGTCCAAGATRAAACGATATAAAGTGTAATGAATTAATTTTATGCTT	1797
RESULT	4			
LOCUS	A67815	4081 bp	DNA	PAT 05-MAY-1999
DEFINITION	Sequence 20 from Patent WO9743427.			
ACCESSION	A67815			
VERSION	A67815.1	GI:4756638		
KEYWORDS				
SOURCE				
ORGANISM	thalia cress.			
	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;			
REFERENCE	Arabidopsis			
AUTHORS	I. (bases 1 to 4081)			
TITLE	De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.			
JOURNAL	PRODUCTION OF APOMICRIFIC SEED Patent: WO 9743427-A 20-NOV-1997; CIBA GEIGY AG (CH)			
FEATURES				
SOURCE	Location/Qualifiers			
	1..4081			
	/organism="Arabidopsis thaliana"			
	/db_xref="taxon:3702"			
	/clone="ARABIDOPSIS SPK GENE"			
	exon	1280..1367		
	exon	1796..1928		
	exon	2014..2085		
	exon	2203..2246		
	exon	2450..2521		
	exon	2617..2688		
	exon	2772..2884		
	exon	3015..3146		
	exon	3305..3646		
	exon	3760..4081		







TITLE Direct Submission  
JOURNAL Submitted (04-JUN-1999) Plant Gene Expression Center, 800 Buchanan  
REFERENCE Street, Albany, CA 94710, USA  
AUTHORS 4 (bases 1 to 88401)  
TITLE Theologis.  
JOURNAL Direct Submission  
COMMENT Submitted (17-JUN-1999) Plant Gene Expression Center, 800 Buchanan  
On Jun 17, 1999 this sequence version replaced gi:4996902.  
The sequence of BAC F23M19 from Arabidopsis thaliana chromosome 1.  
FEATURES  
source location/Qualifiers  
1. .88401  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F23M19"  
complement(796..3891)  
/gene="F23M19.1"  
complement(join(796..1116,1205..1271,1349..1432,  
1506..1636,1794..1850,1999..2047,2166..2251,2468..2545,  
2771..2893,2990..3079,3154..3254,3332..3393,3467..3556,  
3638..3891))  
/note="F23M19.1"  
/note="Contains a PF00561 alpha/beta hydrolase fold  
domain."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39602.1"  
/db\_xref="GI:5091614"  
/translation="MEEDILISPSYDILFQALILPIRIYLLALFVLCITFEYNY  
EHEFLGPAILRYFRCAVNLINPSPSLHYGVROCIHLGRVATPMASPIQICFL  
NHEGLPVTYRQLFLISDGLIALDMLTNSDVIDGLSHKSEITKEDTIVIVP  
GLTSSSSATKLALADPAKTCGNVYISNHRGLGCVTSQCFYNACTEDTIRVLDY  
LOHRYPAPIPAIGTISIGANVAVKYLGESEKIPLEGAVATCSFMDLIGRCIRLL  
KOKLYDALITGLQTAQLHPQLRLAWEGIKRSRSTRDNDHATCLVGFETVPL  
YKRSSTOYGVAVAPLCLISALDPLCTKRAIPWDECRANKNIVATLHGHGLAF  
FEGILTGSSLMWRAINIEFLGLVSCSPYMIQRIIVDKSSSGSQEIPSSINQPYLIE  
DGLVAAYKEDQENKTKLQKRGKRPEDVTKRSFKELCKHOTKOSVWMLGYIGMTSP  
LLGLMNYIFRRKKORPTTSTS"  
complement(5693..6732)  
/gene="F23M19.2"  
complement(join(5693..6108,6233..6428,6499..6732))  
/note="F23M19.2"  
/note="Similar to gb|X98322 peroxidase (pxr10) from  
Arabidopsis thaliana."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39603.1"  
/db\_xref="GI:5091615"  
/translation="MRAITASFIFCFGLVPSVFGKIRRGFTDSSCPLESTVSRVAVK  
HHSLNQETVALLMQFHCFFVNRKLLLCVYVGGDASLLDSTPERPSEKSAEAY  
SVRGFEIDEKKELEIVCPKVSQADIVTLAKDISIALAGLANTNVVALIGGHSY  
GVACSLFQDRLKDPRKMSCLKAKLQNTCRGPDNVYDQATPLEVNOYIKQKSO  
RGLIRIQNLGLDOSTSRIVSNFALNMTLFGERFAAMQINGEIVLVGNSEIFTC  
IVVSH"  
complement(10781..14114)  
/gene="F23M19.3"  
complement(join(10781..11318,11416..11460,11670..11746,  
11875..11991,12207..12310,12443..12512,12614..12717,  
12848..12941,13174..13213,13394..13719,13800..14114))  
/gene="F23M19.3"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F23M19.3"  
/protein\_id="AAD39604.1"  
/db\_xref="GI:5091616"  
/translation="MGGLSSRSSVNNAPGQFAHYNGHLLNNNASDLSNHSGESGLK  
DDSPVTEVDNKHSTSESPFVSSGSPONIEDGIRLSRVLSOKSRSSGSOAA  
VAKYSEVSLIGRAGTGLKAVDIVLTGSSMTNLTSGGSSATYVAKGNTSLTSP  
EVANTIVKGANLHMSLSKDSITHLKEVLPSEVQULISKDDELIRIARAADKQHD  
FVTLVTVQRDLVHEHLALDRFEDDYQKIOEENSPSTAQGVGDTLAIIIRTELKSO  
KRVNLKRSKMSIRLIEEMERLVVYHFLLETHFAFGADPDKPAMPPIHKKRLGS

gene  
CDS  
/gene="F23M19.4"  
join(23307..23355,23431..23627,23656..23751,23850..23906,  
24002..24145,24225..24309,24401..24491,24571..24735,  
24842..24955,25045..25117,25210..25348,25453..25865,  
25953..26140,26231..26279)  
/gene="F23M19.4"  
/note="Similar to gb|AF082176 auxin response factor 9 from  
Arabidopsis thaliana."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39615.1"  
/db\_xref="GI:5091627"  
/translation="MSGNVVNNAPLGGITIGSSKSYVEYOIMKICAGPLDIPGCE  
KYTFPGHIEVSSLSLPLPSFLHFSLSLSLSTVETRELBELDIPICDPS  
KLQGRVLAHLKVENNSDEYATILMPDTQVVLPTQENQFRLVNSFTKRLTASD  
TSAGGEFVPRKKAIECLPSLDMSQPLAQBELALDLHGONRFRHNSRGTPORLLA  
TGNNAFTSKVLKAGDIVVREGTESELKAVIRARHQGNIPSSIVSIDCMRQGVVA  
SAKAFDNOCKEYVYVYKPRSSKEIVSGEIDLVANKEFVGSFRTMLEGDEFSERK  
EGTIIGVSDSPHMKCEMSRLWQDEFSFGPKKVSFMDIRLEALINPAVRSFL  
KNKRLREYNETGSSSHLPLPITLQGEENQSLASDPNMSIRYDATEDAMPKIL  
MSYVQPMKNTNNQVTEMEENITTKGTNFRAGVLTDPVYIKQDIEIGSEIS  
KLTEGKFKLSQTLRSFTEIQNKQFSSSRCTKVQKQVIGRAVDSLNGIDDLIL  
ELEKLFIDKGLQTRNQWELAFITDSDDKLVGDDEWPFCKMVKKIFIKRR"  
complement(28888..31377)  
/gene="F23M19.5"  
complement(28888..31377)  
/gene="F23M19.5"  
/note="contains similarity to gi|47935 protein kinase PK1  
from Zea mays, is a member of the PF100954 S-1000  
glycoprotein family and contains a PF10069 Eukaryotic  
protein kinase domain."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39605.1"  
/db\_xref="GI:5091617"  
/translation="MAVKTPIKLIPLLILLHFPSPSTIGSVIYASGNOWPS  
PNSTSVSFPSPNSFLAVSFAVSPIWMSGTYDSRGLRLTSGSLRTNRSGT  
ITWDSKIDRLGVTSGIEDTEFLNNRSPVWSSFDNPTDITVQSGVSTGKILRS  
GLYSPOLRESGNTLRLNNTSAIYWNHNSFSSNSRSLSLQNGVSTJESMLG  
GAEIYSGDYDSTNFRFLKLDDGDKLRIYSRSGSVNNAHWSVQCLVYGCN  
FGICSYNDTNFICSPSRNEDFVNDNRKCRKYELEDGSGNTMLDLVTRLETY  
HDDPSESSEFFAGSSPCRANCLSVILCLASVMSDGNOMOHGSGFEFGVOWPVS  
TSYVYGVPAVNTLRATRGDNNNSVHIVAVVINGILGLVAVETGLMCCCR  
NPRGTSISHTLLEIYASCAPVOFTTELOKCTSKREKELGAGGVTYRGVLTNRYR  
VAVKOLEIEGEEKQFMEVATVISTHMLNLVLIOFCGQGRRLVIEFMKGLDN  
PLFTGDSAKFLTWBYRNIALAGITTYLHECRNIC IYACIKRENLIVDNFLAKY  
SLFGLTAKLLNDRNRYNMSVYRGRIYHEMLANPIITCSIDVSYGYVGLVLEVSGK  
RNFVDESEKTNKRKFSIYAEIEFGKNTKALIDRLIEDDTVMEDQMRVAKSFQICQ  
DQPLQRPMTGVQVQMLGIEIKRNPDKTILSPVSVSGMSMYSNASHMSVAVSGPSS  
SPSARSPQMGITSSGSPASRTISEGMLGS"  
complement(36027..36827)  
/gene="F23M19.6"  
complement(36027..36827)  
/note="Contains two PF100560 leucine-rich repeat domains."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39606.1"  
/db\_xref="GI:5091618"  
/translation="MINYRHLVFCVAVVYVDSRLPYLAIEQVDPYIVKIVLPYGR  
FDPEEFVSMOGNPCFEGNCLGEGIIIGISISLNLGITSPPHADLTSRLVIDLS  
HNRKCTIPETITLKNLITVDYSVNLGHEVYRQYILREPRNISTGLVPSLP  
TNKKRPTVILVILGIIVGVVAGSFGFYLRIRKQRPLOEPEAVTLTQQQSSD  
ESLVDESEVYSIQIQRYLRRFWSKGLLLTLRLQKTNQNHLPYM"  
40313..42256  
/gene="F23M19.7"  
join(40313..40937,41137..41393,41705..42256)



```

/gene="F23M19.7"
/codon_start=1
/evidence=not_experimental
/product="F23M19.7"
/protein_id="AAD39607.1"
/db_xref="GI:5091619"
/translation="MVPKIASMASRPLCSIPISPLSFLLFVVLFFESNLSINP
NPSISHTLONGINVVAELPRSLYGLIDYWSSTPDSIPDPDPPTKRTSPK
VPPENPLIKOYSAEMIMDLESPKRGSPAKRPSFSDADYVPPFALISAE
MELNGKSPFRKSGNEDYQKROYLDVKTATKAKRNGSDHVFVLTDPVAMHVE
ETALSILVVDGNEKODSSNGSTSLPERIQVSVINDVTPYTHLPRDLISQ
NQRHSILYFKAKRHRHGILREKLMDLVNEPVMVEEPNATGEOSIRMSRP
EFCILHAPGPTSCRLFDALDISICIPVTSDTIELFEDGIIIDYSEFVSASVSLP
KWLANHLGFSEREKETRSRAKYOSVFYVNDGNHADIIGPIENGVANHMKVQOK
VPMVEAVIERRERKPGASVPLRCOCI"
complement(43028..48860)
/gene="F23M19.8"
/complement(43028..43288,43400..43672,43778..45067,
45156..45311,45523..45649,45915..46249,46361..47119,
47284..47637,47725..47869,48023..48613,48703..48736,
48815..48860)
/gene="F23M19.8"
/Note="Contains similarity to g1|36774_FAB1 protein from
Saccharomyces cerevisiae genome gblD50617."
/codon_start=1
/evidence=not_experimental
/protein_id="AAD39608.1"

Query Match      5.7%; Score 383.2; DB 8; Length 88401;
Best Local Similarity 65.8%; Pred. No. 1.5e-56;
Matches 625; Conservative 0; Mismatches 258; Indels 67; Gaps 2;

QY 5732 ttattcaagttgtcgtcctcgaattttgacgtacatgtgagctgttca 5791
DB 72330 TTTGTTTTGATCTTATCCATTTTAAACTTCCATACATGATGAGCTTTCTGA 72871
QY 5792 tcaatgtagaagcgtcgaatcagaacctccctgattgcccacatgagagagatt 5851
DB 72870 ATGATTACAGAACGTCACACCTCACAGTTCCTCTAGCCTGGCAGATAGAGCAATTC 72811
QY 5852 gcaactagatccttcaggagccatctaattgacgacatgtgacccaagattac 5911
DB 72810 GCGCTAGAGTACAGAGAGGGTTGTCTTATCTTCAATGATCATTTGCGACCCCAAAATATT 72751
QY 5912 catcgcgatgaaagctgcgaatataatattgacgaaagaaattgagagctgtgagt 5971
DB 72750 CACCGTATGTGAAGCTGCTATATTTGTTGTGACGAGGAATTTGAGCGGAGTGAAGT 72691
QY 5972 gatttgggttaagctgacgtcgaatgataagataccatgttaagctgtgaag 6031
DB 72690 GATTTCGGGTACTGACTATGACTATTAAGACTATGCTCAACAGGCTGTGCGT 72631
QY 6032 gtaaccatgtggacataagctccgagatccctcgcagcggaaagcatcagagaagacc 6091
DB 72630 GGGGCTATTGGACACATGTGCTGAGTACTCTCAACTGGAAGAACTTCAAGAGAAACT 72571
QY 6092 gatgtccttgggttaagggataatgtcctcgaagctcaattactgagcagagggctttgat 6151
DB 72570 GATGTTTTCGCTACGGGATCATGCTTTTGGAACTGATTAACAGAGTCAGAGACTTTTGAT 72511
QY 6152 ctgtgcctcgtcgaagcgtgtagtgtattgtgtgtgattgggtatggtcccgagt 6211
DB 72510 CTTCGACAACTGCGGATGCGATGAGTATGCTCTCTAGATGGGTA----- 72462
QY 6212 gtcccttggttaattacataattagtgactactacttggttgcccttggtt 6271
DB 72462 -----TAAACAGAGTCTT 72430
QY 6272 ttatttcctgctgatttattcctagtcattgacatattgacctgtcttgcacatg 6331
DB 72449 TTACACATATCTGCGCATCTCAAAAGCTGATTTATCTGTTCAATTTGGCT----- 72395
QY 6332 tcctttagtaaaagcctttgaaagagaagaagcttgagatgctgctgcgtccctgacc 6391

```

```

DB 72395 --TCTCAGGTGAAGGGCTTTTGAAGAGACAGACTGGAGTGTCTTGAGTCTTGACC 72338
QY 6392 tgcagaacaattacattgacacagaagattgagcagcttattcaagtagcatctctga 6451
DB 72337 TGCAAAGCATATTACACAGAACGAGAGTGAACAGCTCATCAATGAGCTCTTCTGCA 72278
QY 6452 cccaaggttcgcaatgagagcgcgccataatgtaagtagtccgaatgcttgaagtg 6511
DB 72277 CACGAGACTCACTATGAGAACGACCTTAAGATGCTGAGGTTTTCGAATGCTTGAAGTG 72218
QY 6512 atgaccttgagaaaagctggagcagatggcaaaagttgaatgcatcacaagacgtag 6571
DB 72217 ACGGTTAGCGGAGAAATGGGAGCACTGCGCAGAAAGTGAATCTTCAGCGCAAGAGTG 72158
QY 6572 aattagctccaatcgaactctgtaattgatactcgaactcgaacataacttgatctt 6631
DB 72157 AGCTCTTCTCACCCACCTCTGACTGATCTTGATTCGACTATATCTTCATGCTTA 72098
QY 6632 ttgaattatcgtgccagataaagacatataaattgaatgaattaa 6681
DB 72097 TGGAGTTGTGTGCTCCAGATTAACGACATTTGATTTGCTTAACAGANA 72048

```

## RESULT 6

```

AC012654 98471 bp DNA PIN 10-DEC-1999
LOCUS AC012654
DEFINITION Arabidopsis thaliana chromosome 1 BAC F14023 sequence, complete
sequence.
ACCESSION AC012654
VERSION AC012654.2 GI:6554462
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

```

## REFERENCE

1 (bases 1 to 98471)  
 Liu, S.X., Lee, J.M., Sakano, H., Yu, G., Jhanveri, A., Lenz, C.,  
 Tortum, M., Chin, C., Chou, J., Choi, E., Gonzalez, A., Howing, B.,  
 Koo, T., Li, J., Liu, A., Pham, P., Vaysberg, M., Altatli, H., Brooks, S.,  
 Buehler, E., Chao, Q., Conn, L., Conway, A. B., Hansen, N.,  
 Johnson-Hopson, C., Khan, S., Kim, C., Lam, B., Nguyen, M., Palm, C.,  
 Shinn, P., Tambunga, G., Davis, R.W., Ecker, J.R., Federspiel, N.A. and  
 Theologis, A.  
 Arabidopsis thaliana chromosome 1 BAC F14023 sequence

## JOURNAL

Unpublished  
 2 (bases 1 to 98471)  
 Theologis, A.  
 Direct Submission  
 Submitted (02-NOV-1999) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 3 (bases 1 to 98471)

## REFERENCE

Theologis, A.  
 Direct Submission  
 Submitted (10-DEC-1999) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 On Dec 10, 1999 this sequence version replaced gi:6151511.

## COMMENT

## FEATURES

## BASE COUNT

## ORIGIN

1..98471  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="F14023"

Query Match 5.7%; Score 378.6; DB 8; Length 98471;  
 Best Local Similarity 58.7%; Pred. No. 8.7e-56;  
 Matches 863; Conservative 0; Mismatches 454; Indels 152; Gaps 6;







/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB38783.1"  
/db\_xref="gi:4490292"  
/translation="MGVEIQSPQSSVYVEQVLAALNPNPEILLPLENRYVNTSOTY  
SLEVENCLRLYQEPPEPMNHIVARILYALAMPDPDSICLFLIPPEVQKPEPK  
SLIVSHLYETGERQQQWDEAKNRHILEANPGEQAQLQAYASHLSLSIQKIPRSVL  
AEAVNNDGASLDKFLTEQVYNSGMIVKEGGSIVLPONEFNHPELKNKTGENVPLEHI  
ARIPILG"  
/number=1  
/complement(2572. .2754)  
exon  
/number=1  
/complement(2755. .2837)  
intron  
/number=2  
/complement(2838. .3031)  
exon  
/number=2  
/complement(3032. .3098)  
intron  
/number=3  
/complement(3099. .3179)  
exon  
/number=3  
/complement(3180. .3230)  
intron  
/number=4  
/complement(3231. .3460)  
exon  
/number=4  
/complement(3461. .3574)  
intron  
/number=5  
/complement(3575. .3667)  
exon  
/number=5  
/complement(3668. .3724)  
intron  
/number=6  
/complement(3725. .3814)  
exon  
/number=6  
/complement(3815. .3928)  
gene  
/number=7  
/number=7  
/gene="F17M5.20"  
/complement(4544. .4687)  
exon  
/number=1  
/complement(join(4544. .4687,4781. .5362,5443. .5712,  
5815. .6093,6423. .6491))  
/note="strong similarity to WD-repeat protein -Daucus  
carota, P1D:g2253631: Contains trp-asp (WD-40) repeats  
signature [VLAIALDHTVYLMDA] [TGGGGGDRITKFMNT]  
[VASAGDETLRFMNV]"  
/codon\_start=1  
/product="WD-repeat protein-like protein"  
/protein\_id="CAB38784.1"  
/db\_xref="gi:4490293"  
/translation="MDAGLNRCPLQEHFLPRKNSYENIDRFIPNRSAMNPDYAFALT  
EERKGDQSAIVSSPSKEATRKQLAETMKNLNRILATRNKPQAVELPSNHSASLH  
QQPKSVKPRRYIPQTSERTLDAPDIVDFYLNLDWGSANVLAIALDHTVYLMDASG  
SISELVITIDEKGPVYSINMAPDGRHVAVGLNNSVQLMDSASNRQLATLGQHSQSV  
GSLAMNHHILTTGGMDGLIINDVRIRSPIVETVYRGHTEVCGKMSGQGLASGCGN  
DNVYHITWDSVASSNSTTQWLHRLHEHTSAVAKALMCPQANL LATGGGGDRITKFM  
NHTTGACLSNDTGSQVCSILSKNERELSHGFTQNLITMTPSVYKAEITLGT  
SRVLYMAQSPDGCYVASAGDETLRFMNVGVPETAKKAAKAVAEPSVNRIR"  
/complement(4688. .4780)  
intron  
/number=1  
/complement(4781. .5362)  
exon  
/number=2  
/complement(5363. .5442)  
intron  
/number=2  
/complement(5443. .5712)  
exon  
/number=3  
/complement(5713. .5814)  
intron  
/number=3  
/complement(5815. .6093)  
exon  
/number=4  
/complement(6094. .6422)  
intron  
/number=4  
/complement(6423. .6491)  
exon  
/number=5  
/number=5  
/gene="F17M5.30"  
/gene="F17M5.30"

exon  
/complement(7856. .7999)  
CDS  
/number=1  
/complement(join(7856. .7999,8104. .8685,8767. .9036,  
9139. .9417,9803. .9901))  
/note="strong similarity to WD-repeat protein -Daucus  
carota, P1D:g2253631: Contains trp-asp (WD-40) repeats  
signature [VLAIALDHTVYLMDA] [TGGGGGDRITKFMNT]  
[VASAGDETLRFMNV]"  
/codon\_start=1  
/product="WD-repeat protein-like protein"  
/protein\_id="CAB38785.1"  
/db\_xref="gi:4490294"  
/translation="MDAGMNTSSHYKTAQRCPLQEHFLPRKNSYENIDRFIPNRSAM  
NPDYAFALTEGKRGKDOTAAVSSPSKEATRKQLATRNKPLNRIATFAPNKPQAVEL  
LPSNHSASLHQPKSVKPRRYIPQTSERTLDAPDIVDFYLNLDWGSANVLAIALDH  
TVYLMDSASTGSELTVIDEKGPVYSINMAPDGRHVAVGLNNSVQLMDSASNRQLR  
TLKGHQS RVGSLAMNHHILTTGGMDGLIINDVRIRSPIVETVYRGHTEVCGKMSG  
SGQQLASGNDNRYVHIMDRSVASSNSTTQWLHRLHEHTSAVAKALMCPQANL LATGG  
GGDRITKFMNHTGACLSNDTGSQVCSILSKNERELSHGFTQNLITMTPSVYKAEITLGT  
SRVLYMAQSPDGCYVASAGDETLRFMNVGVPETAKKAAKAVAEPSVNRIR"  
/complement(8000. .8103)  
intron  
/number=1  
/complement(8104. .8685)  
exon  
/number=2  
/complement(8686. .8766)  
intron  
/number=2  
/complement(8767. .9036)  
exon  
/number=3  
/complement(9037. .9138)  
intron  
/number=3  
/complement(9139. .9417)  
exon  
/number=4  
/complement(9418. .9802)  
intron  
/number=4  
/complement(9803. .9901)  
exon  
/number=5  
/complement(10668. .10967)  
intron  
/number=1  
/complement(10668. .10967,11056. .11159,11587. .12457,  
12511. .12621))  
/gene="F17M5.40"  
/complement(join(10668. .10967,11056. .11159,11587. .12457,  
12511. .12621))  
CDS  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAB38786.1"  
/db\_xref="gi:4490295"  
/translation="MEENCDECMKREBELYWTHTPOTLHFTOLLPGFHNRLQPMFLFO  
FLHYTFICFPAREHLEHTYTQTKRGLSNFCFFLSKSVLCGERSPIANSSITITFC  
NLLOVITPKRFSTHCKRKLPQVTLTKSPSGVAYNGVVEDEKTMARFEGDKVXDS  
LEENDLVFKHGVSESEVLEVDOTLCEKPTSYFKCGHAKRTGIIIDFNATSSRS  
PEHNDIPADVEPTNQOQLVISPVDNLELDLIDLDIDDKILNPLVASHGYROE  
EIRHINDITQOQLVISPSTVAVSEKTYPLSGCKMRRELNDNIDQAAASNRKA  
LSIARASISPGFLVEMKRSVYKCFLLTIPKCYKNNMLITREDEVVMDQDKKMKR  
NIRGARSGGCSITGMKTFVQDNMLREGDVCVPEPANSERKPLHINVYIFRGETERT  
NNVDVYITISE"  
/complement(10968. .11055)  
intron  
/number=1  
/complement(11056. .11159)  
exon  
/number=2  
/complement(11160. .11586)  
intron  
/number=2  
/complement(11587. .12457)  
exon  
/number=3  
/complement(12458. .12510)  
intron  
/number=3  
/complement(12511. .12621)  
exon  
/number=4  
/number=4  
/gene="F17M5.50"  
/gene="F17M5.50"  
/gene="F17M5.50"  
/note="strong similarity to Arabidopsis thaliana







FEATURES		Location/Qualifiers
source		1..65899 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="2" 1..65698
misc_feature		/note="Sequence from clone F13J11" complement(join(451..842,907..1234)) /gene="At2g13660"
mRNA		complement(451..1234) /gene="At2g13660"
gene		/note="F13J11.1" complement(join(517..842,907..1114)) /gene="At2g13660"
CDS		/note="unknown protein" /codon_start=1 /protein_id="AAD28310.1" /db_xref="GI:4726110"
repeat_region		/translation="MCCGRLREKGFANTDSSPEPTFISCHRLITGNTLHOPADOL SPRNLYLCLRLYLAFCHRSPEETEDSDNSSTVPIKINILINPRKDALNKF WMSKELPSLGRRIKHSNMKDPDFWNHFLPEIGRYNLSLOKATRRKRKYVNPKSR ETGIKGNLKHKSSD" complement(2606..2687) /rpt_family="GAA)n" join(<3349..3541,3644..3744,4109..4230,4375..4471, 4583..4693,4792..4928,5629..5755,5848..5997) /gene="At2g13670"
gene		/gene="At2g13670" /note="F13J11.2; predicted by genscan and genefinder" join(3349..3541,3644..3744,4109..4230,4375..4471, 4583..4693,4792..4928,5629..5755,5848..5997) /gene="At2g13670"
CDS		/note="hypothetical protein" /codon_start=1 /protein_id="AAD28311.1" /db_xref="GI:4726111"
gene		/translation="MAOSTSHDSGPOGLMRRRPSRAATVTSIEVDEHYVAVSLGTI APLRYAAIEHEHPRVATLCRYFAEFAEKAHRIDPSGGVGRFOFKTLIFORLDMAN SLASRYKTDGREVEFYOOYEHYVRLADGDDAPROIGKAYCTAGVLEPTMNS KSEKVAVAPETIIAARDYQENETIYAPNITIPDSASQSVQLEEVYKAAVALN TRGLNIPSGEBOHRTKGMJLDLWLRAFEGQASWLPQAAODIQORILTWGLYLIL WGEAANIIRMPBCLCYIFHNMAVELHGLAGNVSIVTGENIKPSYGGDEAFLEKVT PIRVYQVQV" 10367..13711 /gene="At2g13680"
mRNA		/note="F13J11.3" /pseudo complement(join(<14929..15843,16401..>17114)) /gene="At2g13690"
gene		complement(<14929..>17114) /gene="At2g13690"
CDS		/note="F13J11.4; predicted by genscan and genefinder" complement(join(114929..15843,16401..17114)) /gene="At2g13690"
		/note="hypothetical protein" /codon_start=1 /protein_id="AAD28312.1" /db_xref="GI:4726112"
		/translation="MGSDSLRKPTNGTGHLSRRRSCSPFAVPAPSDPTSRSNH IPAKSOQREKLVPSCPOSSKALNVNIDRRLISGVRSPIDSDPTVMDETF TOBEDDAVYVDSPTPLRSRSPAKITVYTGSLSEGVDARSLKGRNGGVILIES LRYLANSDVPSGLIEKRCSSSSSLKNTCTIETVDENLQVFEKRYGLMRES NVILKRMVGTVAIDVLEVAGIFSRVAVLSCTVLEAVPWTEDEKIRRLGIT SFDDAVSELTIAFNSNETENLQDSLTKLTVSITSCSDVNPRLKSLVGLCKSS VYKEQPEIKEDIVRAGKCVDSLAKLEBSSSSSKKEKPLTSLRVEVNNML LEIMIDIEAEFEVIMGKORRLVHEHNSVYEVSRVYGAFLIAGRKRYVCCG EARAGVIAEAFVPMVDFGMLQCKKGLDMREVEGMOGTLITLVKEQYVEMEMER WFSKHGTECPNLKAFQIWMRSRFLGVSSTCR"
gene		complement(18082..20371) /gene="At2g13700"
		/note="F13J11.5"

repeat_region	/pseudo 29450..29498 /rpt_family="POLY_A"
gene	complement(29932..30854) /gene="At2g13710"
	/note="F13J11.6"
repeat_region	/pseudo 31008..31094 /rpt_family="GAA)n"
repeat_region	complement(31832..31913) /rpt_family="GAA)n"
mRNA	join(32232..32373,32688..32771,32839..32939,33013..33167, 33252..>33473) /gene="At2g13720"
gene	32232..>33473 /gene="At2g13720"
	/note="F13J11.7"
CDS	join(32264..32373,32688..32771,32839..32939,33013..33167, 33252..33473) /gene="At2g13720"
	/note="unknown protein" /codon_start=1 /protein_id="AAD28313.1"
	/db_xref="GI:4726113"
repeat_region	/translation="MVALVAOERLGSMRCSMRDPRGRMPRVPRVPRDLKSPITV LVTPYKVSRLTSSSHRYFTVLTSTVSTLWTRGSKKDRNYLTKRYKGVV QVREKNTLSVNDONKRFITDSSSEALNAGTSDOKRWETKVTGFTLDSKKQ QIMYSDFTKRLKLTFTTSNQRHININKNYILKNDKSKNIIILKKYKVVLYII IVLIL"
repeat_region	33343..33441 /rpt_family="POLY_A"
mRNA	complement(<33782..>34228) /gene="At2g13730"
gene	complement(<33782..>34228) /gene="At2g13730"
CDS	/note="F13J11.8" complement(33782..34228) /gene="At2g13730"
	/note="unknown protein" /codon_start=1 /protein_id="AAD28314.1"
	/db_xref="GI:4726114"
	/translation="WADVIRADAKGYRVEDGOAVNEAQRIDHALIPEDANDQ AALNQLAARDAAGLIVLDLNRGIDRLHWNQNDALGVENHNNGVRHPOIPAAITL AAEPRIITGDFNRPGMFANRSALVPPPYORNDYESHLYITLHIT"
repeat_region	complement(35689..35776) /rpt_family="TAAAA)n"
gene	36924..37497 /gene="At2g13740"
	/note="F13J11.9"
repeat_region	/pseudo complement(37457..37522) /rpt_family="GAAA)n"
mRNA	join(<38939..39051,39561..39686,39797..39989, 40095..>40208) /gene="At2g13750"
gene	<38993..>40208

Query Match 4.7%; Score 313.6; DB 50; Length 65899;  
Best Local Similarity 64.8%; Pred. No. 1,2e-44;  
Matches 542; Conservative 0; Mismatches 239; Indels 55; Gaps 3;

QY	5716	taatgtcccaacttaattcaatgttagagcgtaagcaatcaagaccctccctgattgccc	5835
DB	64398	TATTGTGCAATATTTTGTGTTTGTGTTGACGTTGTGATTTGCTGAGTTGTAAGTAT	64457
QY	5776	tgatgtctgttttcataatgttagagcgtaagcaatcaagaccctccctgattgccc	5835
DB	64458	CTTATCTTTGATTT--TCATATATAGAGGCTCTGAAGCAATCAGCCCTTGACTGCC	64515
QY	5836	aactaaggagaagatgcactagtagctctctagggccattatcaattgcattgacattg	5885
DB	64516	AAAAAGAAACATATTGCTCTGGGATCAACAAGGGGCTCGCATTTTACAGATCATTG	64575







OY	6580	ccacatcgcgaacttcgaaagtgctcctagatcgcgcgaagataacttgcattgtttgatcta	6639
Db	1998	CTTAATCCCTACTCIGATTGGATTTCTTAAITCTACTTACAAATTTCGACGCCGTAGATTA	2057
OY	6640	tctggtccaagaataacagcatataaaa	6667
Db	2058	TCTGGTCCAAGGATAAAAAAAAAAAAAA	2085
RESULT	10		
LOCUS	AB013395/c		
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MON23, complete sequence.		
ACCESSION	AB013395		
VERSION	AB013395.1	GI:3128142	
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsut P1 clone:MON23.		
ORGANISM	Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;			
REFERENCE	Arabiidopsis. 1 (sites)		
AUTHORS	Nakamura,Y.		
TITLE	Structural Analysis of Arabidopsis thaliana Chromosome 5. VI		
JOURNAL	Unpublished (1998)		
REFERENCE	2 (bases 1 to 86064)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-MAY-1998) to the DDBJ/EMBL/GenBank databases.		
FEATURES	Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2; 1532-3, Yana, Kisarazu, Chiba 292, Japan (E-mail:Ynakamu@kazusa.or.jp, Tel:+81-438-52-3935, Fax:+81-438-52-3934)		
source	Location/Qualifiers 1..86064 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="MON23"		
BASE COUNT	27110 a 15313 c 14986 g 28055 t		
ORIGIN			
Query Match	3.2%; Score 217; DB 7; Length 86064; Best Local Similarity 55.6%; Pred.No.3.9e-26;		
Matches 513; Conservative 0; Mismatches 335; Indels 75; Gaps 2			
OY	5751	cctcgaaatttgatcgtfacattgfatgctctgttttc--atcaaatgtagcgcgc	5807
Db	73777	CTATAAACCTTTTGATTTCAGAGAGTTTAAGAATTTGTGATCTATATPACTATGAACAGAGATT	73718
OY	5808	aggccatagaacctccccctgattggccaactagygagagagattgcactggatctctta	5867
Db	73717	AACCCGGGGGCACAGTTCGATTGGTTTCAGGAGGAACAACATTTGCCGTAGGTGCACAC	73658
OY	5868	ggggcccatthaattgcataaccattgcatgcccaagaattatccatgcgcattgaag	5927
Db	73657	GAGGACCTCGAATNTCTTCATCAAGACATTCGAACCCGAAGATCATCACAGAGATGTGAAG	73598
OY	5928	ctgcaatatattattgagcagaagtggagagctgttgttagtgatttgggttagcta	5987
Db	73597	CTGCAAAATGATTACTACAGATGAGAGACTTTTGAAGCAGTGGTGGGATTTTGTTTAGCCA	73538
OY	5988	ggctcagtggttacaaagatcacccatggttagcagctgctgtaagggtgacattgggaca	6047
Db	73537	AGTTGGTAGATGTTAAGAAGACTAATAGTAAACCACTCAGGTCGAGGAACAATGGGTGATA	73478
OY	6048	tagctcccggtaacctctgcactggaagtcatacagagaagaccgattgtcttggtaatg	6107

[illegible]



```

/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="WML2"
/clone.lib="Mitsui PJ"
BASE COUNT      27260 a 15039 c 15617 g 26980 t
ORIGIN

Query Match      3.1%; Score 208.2; DB 7; Length 84896;
Best Local Similarity 59.9%; Pred. No. 1.3e-26;
Matches 443; Conservative 0; Mismatches 238; Indels 53; Gaps 3;

QY 5811 catcaaacctccctcgtattgccaactaggaagaggtgtcactagatcttctaggg 5870
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67256 CACCTAAACCGGTATGATGGGSCACAAAGAGGATACATTAGAGAGAGAGAG 67197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5871 gctatcataattgcatgccaattgtatcccaagattatcatcgagatgcaaaagctg 5930
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67196 GGTTCGTGTTTGCATGAGCATGTGATCCAAAGATCATTCACCGTGTGTCAAAGCTG 67137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5931 caaatatattattgagcagaagaattgagcgctgtgtagtgatattgggttagctagcg 5990
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67136 CGAACATACTTCTTGACGATTACTTGAAGCTGTGTCGAGATTTCCGGGTGCTTAACG 67077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5991 tcattgattcaagagatcccatgtagacgtctgtaagggttaccattgggacatag 6050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67076 TTTTGGATCATGAGAGTCCATGTGACACACCCGCTGAGAGAGACAGTGGGTCACATTTG 67017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6051 ctccagatcctcctcagctcgtgaaagtcacagagaagaccgactgtcttggtttaggga 6110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67016 CACGTGATGATCTCTCAACAGGACAACTCTCTGAGAAAGACATGTGTCGGGTTCGGGA 66957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6111 taatgtctccagagctcatcactcgtgacagagagggcttctatctgtcgtccttgagaag 6170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66956 TTCTTCTCTCGAATTGATTACTGATGAGAGCTCTTGATTCGGAAAAAGCAGCAAAAC 66897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6171 atgatgattcatgtgtgattgattggtatgtgtccgggtgtccttggtaattatt 6230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66896 AAGAGAGAGCGATCTT---GATTGGGTAG-----AAG 66866
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6231 tcacatatagtgcttactactctgtgtgtgcccccttgttttatttccctgcctgattt 6290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66865 TAAAGAGGTAAACATTTCTCTCTGTTAGTTAGTTAAATAATCTCACTCTCTAG 66806
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6291 gattcttagtcagtattatgatatgacccgtcttgcattgcttctttaggttaaaagcct 6350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66805 GTTTTTT-----CTTTGGGGTTACTTGTAGGTAAAGAACT 66770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6351 tttagaagagaaaaagttagagatgctgtgcagctccctgacctgcaagaacattacattga 6410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66769 ACAACAAGAGAGAGAGCTTAACAGATAGACAGCAAGATTTAAGAAGCAGACATG 66710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6411 cacagaagttgagcagcttattcaagtagcattactctgtaccagaggttcgccaattga 6470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66709 AATGAGAGTGGAGAAATGCTCAAGGCTTTGCTTGAACACATATCTTCCATCTCA 66650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6471 ggcggcgaagatgtcagagtagtcgaaatgcttgaaagtgatggccttgacgaagaagtg 6530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66649 CCGGCCCTAAGATGTCTGAAGTGTGAGAAATGCTTGAAGGCGGATGCTTGTGAGAAAAG 66590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6531 ggaagc 6535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66589 GGAAG 66585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
AC002292/c AC002292 120787 bp DNA PLN 02-OCT-1997
LOCUS AC002292 Genomic sequence of Arabidopsis BAC F8A5, complete sequence.
ACCESSION AC002292
VERSION AC002292.1 GI:2252639

```

```

KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsids.
1 (bases 1 to 120787)
REFERENCE
AUTHORS
Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,
Kim,C., Li,Y., Ojima,O., Osborne,B., Shinn,P., Sun,H., Toriumi,M.,
Vysotskaya,V.S., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Genomic sequence of Arabidopsis
Unpublished (1997)
2 (bases 1 to 120787)
REFERENCE
AUTHORS
Federspiel,N.A., Davis,R.W., Conway,A.B., Palm,C.J., Conway,A.R.,
Kurtz,D.B., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Shinn,P.,
Sun,H., Ojima,O., Osborne,B., Shinn,P., Toriumi,M., Vysotskaya,V.,
Yu,G., Theologis,A. and Ecker,J.
Direct Submission
Submitted (05-JUN-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
3 (bases 1 to 120787)
REFERENCE
AUTHORS
Federspiel,N.A., Davis,R.W., Conway,A.B., Palm,C.J., Conway,A.R.,
Kurtz,D.B., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Shinn,P.,
Sun,H., Ojima,O., Osborne,B., Shinn,P., Toriumi,M., Vysotskaya,V.,
Yu,G., Theologis,A. and Ecker,J.
Direct Submission
Submitted (12-JUN-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
4 (bases 1 to 120787)
REFERENCE
AUTHORS
Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,
Kim,C., Li,Y., Ojima,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M.,
Vysotskaya,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (09-JUL-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
5 (bases 1 to 120787)
REFERENCE
AUTHORS
Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,
Kim,C., Li,Y., Ojima,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M.,
Vysotskaya,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (11-JUL-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
6 (bases 1 to 120787)
REFERENCE
AUTHORS
Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
Ojima,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vysotskaya,V.,
Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (02-OCT-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
On Jul 11, 1997 this sequence version replaced gi:2251217.
COMMENT
FEATURES
source
gene
CDS

```

```

1. /organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
177. 2876
/gene="F8A5.1"
.join(177. 214.632. 707.1132. 1152.1283. 1479.1617. 1761,
1845. 2159.2256. 2526.2788. 2876)
/gene="F8A5.1"
/note="Hypothetical protein"
/codon_start=1
/protein_id="FAB71954.1"

```



gene		<p>/db_xref="GI:2462735" /translation="MCCRNVTLRQNDPSTEVSLSYGSGMDVLPDIIIGFQKIIYVLOI LVNLTDLNWKQTSISPGNQIYFAVNAKTCPTASNLERKSGLEIYVLAHACIDLT MDCYCFSDREHKSGVQYQCDKHKRYGGMETVYISDLSESTOCRSKCNKTEV LYENDELPSVPHALISALKIDMKYGIILANVODGDHGLFEMDPEISVYLOIL HMYNHPYPRKNGPGISLKKIGIKNALDINKAKHEGFLSSHRKICSPYDARSI AULIISSTIDIDROGDCISLYGQTOEYERDVENTQKITYTVGNMERKPOKNOEA PFLFDGESEITFEFDEYEDENSTHEIDATKATYQFNCLATMOE" complement(3171..4444) /gene="F8A5.2" complement(join(3171..3389,3471..3605,3706..4029, 4118..4444)) /note="F8A5.2" /note="nearly identical to rice water stress induced protein gpID265371537404" /codon_start=1 /protein_id="AAB71970.1" /db_xref="GI:2462751" /translation="MAPEISVNPMTLSEKAOAPPRAVYVTELAGNDPYKGVGLAK GLRKYASAVPLVAMLPDYREHREILSOGCVREIEPYVPDNOYFEAMAYVLYK SKLRINFEESKMIYLDNDIOVPNDIHLTDLSDATFYAMDCCECTKNSHQYSL GTQOCPEKVTWPEDESPPLVIFNAGHVFEPSPLTYSLLDTLETTPSPYAEOD FLNMEFEKYKPIPLVYNLVAMLRHHPENVELEKYVYHCAAGSKRWYTGSEAN DREDIKMLVDKWDVYNDSEISDFKSKIPADAEIVTKSSIIASVLEPMTYFPAPSA "</p>
CDS		<p>complement(8347..9066) /gene="F8A5.3" complement(join(8347..8443,8539..8657,8941..9066)) /gene="F8A5.3" /codon_start=1 /product="putative ADP-ribosylation factor" /protein_id="AAB71955.1" /db_xref="GI:2462736" /translation="MGLSPKLPSPRLFAKEMRIIIVGDAAGKTIILYKIKELIDE LRNAVLYFANKODLPNANMAEITDKGLSLRQRHWTIQITCATTEGLEGIDML SNRIANKVSTF" complement(9782..14180) /gene="F8A5.4" /note="F8A5.4" complement(join(9782..9919,10031..10178,10275..10418, 10509..10663,10780..10917,10991..11164,11286..11477, 11815..11936,12047..12182,12300..12395,12475..12622, 12803..12981,13104..13313,13414..13572,13659..13753, 13867..13963,14069..14180)) /gene="F8A5.4" /note="location of EST gb W43464" /codon_start=1 /product="phosphatidylinositol 3-kinase" /protein_id="AAB71971.1" /db_xref="GI:2462752" /translation="MGANERFELSCDINSPTFRLEKIDGMLPYKSSDSGVSIAT EKKPELYIECALIDAGPGLPWRTRKTTGTPYWNELITLSSYRDLTAHSAIAT VWDVSCGTGALIGATVGLFNSKQMGKGRKLMQKEDSGFPTSTGKVRBE RGLEHLEKLMNKYERGOIOSIDMLRMLKSLDITKDESKHSGSLFVYIDCSF EHRVYRQESANLFTIAPRGSTNEPYVVDLTGTPNEKOLATSLIDGIDRD LKRSNIBKRSIOYVLYKPTPTILSGDERDLMKFRSLSEKRLTKLRCEYSDVO EAKQAIQLMKWMDVCDALTELSPLESEEVRAVASLERADEDELQCLQLQVQ ALAFERSDRCSLSQFLVORALONIELASFLRYVAVELDHVATRFSTYELLENI IKLPYVNGEDGYLQWOSLYRQTELTALQCSITREVRNVRGTQKRIKRLQGLGL SELTYEERIPRSPLTNPVLIKIVAGESSLSFKASHPLLTLRTPEEGSSCLIKKG DLRODOLVOWYMWLMDRLKLENDLCLTPYKVALYDEDEMLFETSRSLAOLISE HRSITSYLOKEHDEHAPGTATACDPTKICACAGSYVYIIGIDGHDLLITDD GRLFHYDFATILDRDPRPPPKKLCKEVEMAGAESQYTRFRSYCEANVIRKS SNLNLNLFHLAGSTIPDIASDERGILKLEKFRIDMDDECIFHFDLLINESVAL PQWVETIHRMAQYMR" 15572..17812 /gene="F8A5.5" join(15572..16004,16236..17812) /gene="F8A5.5" /note="similar to "Mx" GTP-binding proteins" /protein_id="AAB71966.1" /db_xref="GI:2462747"</p>
gene		<p>/translation="MYKKVATKKNSPSLIAKKKRSNDVYSVAPITTSYNDRI PLDITVDRLNINWREGEIHLPITIVVGDSSGKSVLESLAGISLPRGGICTRPL VNRLORSSPEPEIMLEYNVVDDEHIAEAIRATYIPASIGVSGAPDILLVAK KAGVLDLMAVLPGLITRVYPNGEENIEQISGKSVLESLAGISLPRGGICTRPL CSISMSKRVKRGORTLAVTKADMAPEGILQAVDYSIVAGYCVNRIGSEET EARNQDELFRTHVLSLIDEDIVGIPVAQKMLIOSMIRACLPKYKINOKD TRVLENLKPMVMASTGSAALMADIIIGSAKESLIRIYNGDSEVPDDNMCTARL AMLSQESDIOAKREYERFLMDEIKILDECKCVGLPNTIPSAFALISQHVQIGQ DKPVEINKIMDYIEDVLSVYAKRSDNFPQIOSIKRAGRNLISKIKEDSVRVEI YVMEKLTIDYCNPEYMTSMQTKSAQSEFIDAVYKNENIPDYSEYVGFNKKISHLK YVHMLIIPAFDMKNITSYKIVLRIDYNTALYLOLSVLTWTRFOKEIVAEWDP RGGGVKMEESPVIAKREKIQNSIKILKESDVAALVQNC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188, 21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSKRVHTRTSSPILATIOANPHNDREYVPIEADITSSYND IRPLDITVDRLNINWREGEIHLPITIVVGDSSGKSVLESLAGISLPRGGICTR PLVNRLORSSPEPEIMLEYNVVDDEHIAEAIRATYIPASIGVSGAPDILLVAK KAGVLDLMAVLPGLITRVYPNGEENIEQISGKSVLESLAGISLPRGGICTRPL CSISMSKRVKRGORTLAVTKADMAPEGILQAVDYSIVAGYCVNRIGSEET EARNQDELFRTHVLSLIDEDIVGIPVAQKMLIOSMIRACLPKYKINOKD TRVLENLKPMVMASTGSAALMADIIIGSAKESLIRIYNGDSEVPDDNMCTARL AMLSQESDIOAKREYERFLMDEIKILDECKCVGLPNTIPSAFALISQHVQIGQ DKPVEINKIMDYIEDVLSVYAKRSDNFPQIOSIKRAGRNLISKIKEDSVRVEI YVMEKLTIDYCNPEYMTSMQTKSAQSEFIDAVYKNENIPDYSEYVGFNKKISHLK YVHMLIIPAFDMKNITSYKIVLRIDYNTALYLOLSVLTWTRFOKEIVAEWDP RGGGVKMEESPVIAKREKIQNSIKILKESDVAALVQNC" 22510..25157 /gene="F8A5.7" join(23787..24210,24676..25157) /gene="F8A5.7" /note="similar to "Mx" GTP binding proteins" /codon_start=1 /protein_id="AAB71957.1" /db_xref="GI:2462738" /translation="MGSKMSNDREIDVEMASLSIYVPIEAPYTSYNDRIPL LTVDRRLNINWREGEIHLPITIVVGDSSGKSVLESLAGISLPRGGICTRPLVNR LORSSPEPEIMLEYNVVDDEHIAEAIRATYIPASIGVSGAPDILLVAK KAGVLDLMAVLPGLITRVYPNGEENIEQISGKSVLESLAGISLPRGGICTRPL CSISMSKRVKRGORTLAVTKADMAPEGILQAVDYSIVAGYCVNRIGSEET EARNQDELFRTHVLSLIDEDIVGIPVAQKMLIOSMIRACLPKYKINOKD TRVLENLKPMVMASTGSAALMADIIIGSAKESLIRIYNGDSEVPDDNMCTARL AMLSQESDIOAKREYERFLMDEIKILDECKCVGLPNTIPSAFALISQHVQIGQ DKPVEINKIMDYIEDVLSVYAKRSDNFPQIOSIKRAGRNLISKIKEDSVRVEI YVMEKLTIDYCNPEYMTSMQTKSAQSEFIDAVYKNENIPDYSEYVGFNKKISHLK YVHMLIIPAFDMKNITSYKIVLRIDYNTALYLOLSVLTWTRFOKEIVAEWDP RGGGVKMEESPVIAKREKIQNSIKILKESDVAALVQNC"</p>
CDS		<p>complement(26476..28975) /gene="F8A5.8" complement(join(26476..26936,27030..28070,28540..28975)) 2..9%; Score 191.6; DB 8; Length 120787; Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
gene		<p>Query Match Best local Similarity 68.2%; Pred. No. 8.2e-24; Matches 266; Conservative 0; Mismatches 124; Indels 0; Gaps 0;</p>
CDS		<p>Query Match Best local Similarity 68.2</p>



Db 108419 TGTGGCCACATTGACCTGAGTACTTATCCAGCGGTCACTGCTCAGAGAGAGTATG 108360  
QY 6097 ctttggtatgagatactctctagagctcattactgacagaggcttttattcttc 6156  
Db 108359 CTTGGCTTTGGCATCTCTCTTGGCTCATTACTGTCAGAAAGCTCTGATTTGG 108300  
QY 6157 tcgccttgacagatgatgtatgtt 6186  
Db 108299 CAGATCCGCACACCAGAAAGTGTATGCT 108270

RESULT 13  
LOCUS F17123 134784 bp DNA PLN 20-JUN-1999  
DEFINITION Arabidopsis thaliana BAC F17123.  
VERSION AF160182 GI:5106766  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsi.  
1 (bases 1 to 134784)  
REFERENCE  
AUTHORS Washington University Genome Sequencing Center.  
TITLE The A. thaliana Genome Sequencing Project  
JOURNAL Unpublished (1997)  
2 (bases 1 to 134784)  
AUTHORS Sun, H., Wohldmann, P., Johnson, D. and Gibson, A.  
TITLE The sequence of A. thaliana F17123  
JOURNAL Unpublished (1999)  
3 (bases 1 to 134784)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUN-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63108, USA  
e-mail: twilson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by  
M. Iodhi, Cold Spring Harbor Laboratories, and fingerprinted  
by M. Marra, Mashu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we provide a small overlap between  
neighboring submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate  
chemistry; an attempt was made to resolve all sequencing problems,  
such as compressions and repeats; all regions were covered by  
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

Actual start of this clone is at base position 1 of F17123; actual  
end is at 134784 of F17123.  
Location/Qualifiers  
1. 134784  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/clone="F17123"  
/chromosome="IV"

BASE COUNT 43094 a 23410 c 23696 g 44584 t  
ORIGIN  
/map="unknown"  
Query Match 2.9%; Score 191.2; DB 8; Length 134784;  
Best Local Similarity 56.7%; Pred. No. 9.5e-24;  
Matches 430; Conservative 0; Mismatches 273; Indels 55; Gaps 2;

QY 5812 atcagaacctccctcgtatgccaactaaggagaagattgcaataagacctcttaagg 5871  
Db 46525 AGCTTAACCGGATTTGACTGGAACATGAGAGAGATACCAATTTGTCGACGAGAGG 46584  
QY 5872 cctatcctaattgcatgcccattgatacccaagattaccatccgcatgaagaagctc 5931  
Db 46585 TTTGTGTATCTACTATGAGCATGTGATGCCAAGATCATTCATAGATGTAAGGAGC 46644  
QY 5932 aaatatattatgacgaagaatttgaagctgttgaagtttgaagtttgaagct 5991  
Db 46645 TAAATATCTCTAGACGAGTGTGTAAGCTTTGTTGTTGACTTTGGACTCGCAAGCT 46704  
QY 5992 catgattacaagaagttaccatgtaagctgttaaggggtaccattgggacatagc 6051  
Db 46705 CCTTAACCATGCGGATTTCTCATATGTCACACATGCGGTCCGTGACGGTTGACATTCG 46764  
QY 6052 tcccgagtaacctcgcagtggaagatcacaagaagaacgcatctctgttattggat 6111  
Db 46765 ACTGTAATATCTCTCCACTGTCAGTCTTTCGAGAAACGATGTGTTGGCTTGGAT 46824  
QY 6112 aatgctcctagagctcattactgacagaagggctttgattcttcgcgcttcgaacga 6171  
Db 46825 ACTATTTGCTCGAGCTCATATACCGGACTGAGAGCTTTGAGTTGTAACCGTTTACCA 46884  
QY 6172 tgatgatgtatg--ttgttgatggatggtatggtcccggtgttccttggttaattat 6229  
Db 46885 GAAAGAGCTATGCTTGATGGTGAAGAAATGACACTGAGAACTTCTTGATTAC 46944  
QY 6230 ttacatatatgcttactacttctgttggcccttgttttattcttcgctgatt 6289  
Db 46945 TTCATAGATCAACCAATTAACATTTTCTCCGATGCTGTGTA----- 46990  
QY 6290 tgattcctatgcatgttatgcatattgacctgcttgcaatgctctttaggttaagcc 6349  
Db 46990 -----ACATTTGGTTTATGGTGAAGAAAT 47011  
QY 6350 ttttgaagagaagaagtgtgagatgctgctgacatccgactgcaagaacattacatg 6409  
Db 47012 TACATGAAGATGAAAGTAGAGGAACATTTGATCGAAGACTCGGAACATACATGAT 47071  
QY 6410 acacagaagttagagagcttattcaagtagcatctctgtaccgaaggttcgcaatg 6469  
Db 47072 AGATTAACTTGAGAGATGTGCAAGTGGCTTCTATGACACAAATATGCGCAGCTC 47131  
QY 6470 agcgagcctaagatgtcagaggtgagtcggaatgcttgaaggggtgagtgcttcagaagaagt 6529  
Db 47132 ATCGTCTCTAAATGCTGGAAGTTGTTTGATGCTTGAAGGCGAGATTTACCGAGAT 47191  
QY 6530 gggacgagtgycacaagaagttagatcatcatcaagac 6567  
Db 47192 GGGCTGCTTCGATTAACCATTCACATTTCTACCATGCC 47229

RESULT 14  
LOCUS AB005234 89779 bp DNA PLN 20-NOV-1999  
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MBK5,  
complete sequence.  
ACCESSION AB005234  
VERSION AB005234.1 GI:2264306  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone: Mitsu1 pl  
ORGANISM Arabidopsis thaliana







repeat\_region /rpt\_family="(TAA)n"  
complement(join(1549, .2058,2731, .3012))  
/gene="At2g23810"  
gene complement(1549, .3012)  
/gene="At2g23810"  
/note="F2714.1"  
CDS complement(join(1744, .2058,2731, .3003))  
/gene="At2g23810"  
/codon\_start=1  
/product="similar to senescence-associated protein"  
/protein\_id="AA18611.1"  
/db\_xref="GI:6598449"  
/translation="METHLLIVRCITVFAPVYVYVNGAGAIIRGKRYKIDYSTWL  
QKRVENGRNMNKRISCLVESKVCSTLEAFVNPVNSFYKEHLIPLAQSCCPSPEC  
FEVYVPTWTNKTGTHTNPDCOTWNAKEKLCFDCQCKAGLIDNVASMKKALIV  
IVFLVFLIIVYSVGCACFRNNKRDDYSRTYGYK"  
2794, .110211  
/note="Sequence from clone T29E15"  
misc\_feature complement(3567, .3677)  
repeat\_region /rpt\_family="(TAAA)n"  
complement(3570, .3641)  
repeat\_region /rpt\_family="(TAAA)n"  
complement(4565, .4612)  
repeat\_region /rpt\_family="(CAA)n"  
6379, .6484  
/rpt\_family="(CAAAA)n"  
join(6480, .6749,7602, .7726,7811, .7848,7929, .8039,  
8152, .8268,8487, .8786)  
/gene="At2g23820"  
/note="T29E15.2"  
CDS join(6480, .6749,7602, .7726,7811, .7848,7929, .8039,  
8152, .8268,8487, .8557)  
/gene="At2g23820"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAC63656.1"  
/db\_xref="GI:3738315"  
/translation="MKNVNRYFRCISLIVFLVPSFTRSHIRPTYSAGASSPNRAH  
CMADSPSGSGSVSPPNVNAAPSSSSSSASAIIDFSLCTRKTTPRAMIKRDV  
KDPESIAHMTWGLMALISDIPGNRDKCMKALVHDIAEIVGDITPSCGISKEE  
KNRRESLEHMKCLLGGGERAKELAEIMREYENSSPEAKVVDKVELILQALEY  
EQGKQTNIGKAMASEIYSRRRKH"  
complement(6654, .6712)  
repeat\_region /rpt\_family="(GAA)n"  
join(9531, .9582,9709, .9779,9866, .9947,10054, .10169,  
10268, .>10396)  
/gene="At2g23830"  
/note="T29E15.3"  
CDS join(9531, .9582,9709, .9779,9866, .9947,10054, .10169,  
10268, .10396)  
/gene="At2g23830"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAC63657.1"  
/db\_xref="GI:3738316"  
/translation="MSNNLEIEPMTLOPEFLKROMCSLYLTKNTDNNVAFYKT  
TNRNVCYBPNGVGLIPKSTCKVLVTMOAKQEPDSMOSEKFMISYLAPOVTAKK  
VTREMFSESHVYVEIKLRTYVCSITTNITSSPRIRRGFTTCCTC"  
complement(11352, .11371)  
repeat\_region /rpt\_family="(TAAA)n"  
15801, .15917  
repeat\_region /rpt\_family="(TAAA)n"  
15853, .15918  
repeat\_region /rpt\_family="(CAAAA)n"  
16051, .16136  
/rpt\_family="(TAAA)n"  
complement(join(<20442, .20537,20617, .20691,20894, .20948,  
21145, .>21485))  
/gene="At2g23840"

gene complement(<20442, .>21485)  
/gene="At2g23840"  
/note="T29E15.4"  
CDS complement(join(20442, .20537,20617, .20691,20894, .20948,  
21145, .21485))  
/gene="At2g23840"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAC63658.1"  
/db\_xref="GI:3738317"  
/translation="MAHFSARLKLISNDGLISFGVDSRDFQSLYVAGGSPVPIR  
MRLEKFCARISFYSFPPYKRRKIRDLNTERKNDLIDEDNDEMDFDGDGLFEDHLS  
CFRGVLIDSYSPVNVCMKRAICLEYMDKADVLEYDQVSSPTGSFYIPAVLRVPH  
LLQVYKRRVRKSLSRKNILRLDDYTQC"  
complement(22086, .22781)  
gene /gene="At2g23850"  
/note="T29E15.5"  
/pseudo join(<22999, .23070,23189, .23296,23402, .>23554)  
/gene="At2g23860"  
/note="T29E15.6"  
gene /gene="At2g23860"  
/note="T29E15.6"  
CDS join(22999, .23070,23189, .23296,23402, .23554)  
/gene="At2g23860"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAC63659.1"  
/db\_xref="GI:3738318"  
/translation="MTNPKNYCRPNGLILPSTCKLRKPKPSMOSNEFMFOR  
VNASPVAITKEVLTLETFNKESGHLVEETKRLRTYVCSITTNITSSPRTNGLYENFV  
RHLEFSL"  
27493, .27571  
repeat\_region /rpt\_family="(TAAA)n"  
complement(29076, .29146)  
repeat\_region /rpt\_family="POLY\_A"  
29302, .29885  
gene /gene="At2g23870"  
/note="T29E15.7"  
/pseudo 31472, .31519  
repeat\_region /rpt\_family="(TA)n"  
complement(join(<31877, .32176,32267, .32407,32498, .32893,  
33162, .35713,35886, .>36147))  
/gene="At2g23880"  
complement(<31877, .>36147)  
gene /gene="At2g23880"  
/note="T29E15.8"  
Query Match 2.68; Score 176; DB 50; Length 110211;  
Best Local Similarity 66.84; Pred. No. 3,9e-21;  
Matches 251; Conservative 0; Mismatches 125; Indels 0; Gaps 0;  
QY 5811 catcgagaccctcccgatggccaactagggaggaagattgcactagatctctaggg 5870  
DB 53910 CAGCTTAACCCAGCGTTGGACCTCGAACAAGAAAGATAGCGATTGCAAGAG 53851  
QY 5871 gcctactcaattgcatagcattgtgatcccaagattacatccatcgagatgaagctg 5930  
DB 53850 GGTGTTTATCTACACAGACGAATGCGATCCCAAGATTATTCACGAGATGCAAGCGAG 53791  
QY 5931 caaatatatttggacgaagaatttgaagctgttgaagtgttttggtagtgaagc 5990  
DB 53790 CAACATTCTCTAGATGAGTATTGTAACACAGATGTTGGGATTTGGACTAGCAAAAC 53731  
QY 5991 tcatgattacaagatacccatctttagactctgttaaggggtacacattgggacatatg 6050  
DB 53730 TACTCAACACAGAGATTCATATGTCAACAACCGGGTTAGAGAACGTGTGGTCACATTG 53671  
QY 6051 ctcccgagctacctcgactggaagtcacatcagaagaagaccgagctgtcttggttatggga 6110  
DB 53670 CACGTGAGTATCTCTCCACCGGTGATCATCTGAGAAACCGATGTCTTGGGTTGCGTA 53611



Tue Jun 27 14:47:57 2000

us-09-180-798-1.rge

Page 21

Qy	6111	taatgcccacaaagccatattgacaagagagcctttgactctgctcgctcgacgaag	6170
Db	53610	TACTTTGCTAGAGCTCATCACAGAAATGAGAGCTCTCGATTGTGGCAAGTCTGTAGCC	53551
Qy	6171	atgctgctgttatgct	6186
Db	53550	AGAAAGAGCTATGCT	53535

Search completed: June 23, 2000, 23:30:33  
Job time: 43188 sec











CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match 99.9%; Score 6686.2; DB 1; Length 6695;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6692; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tcgaatgacgaatcgccctaccccttgatttngaaatactagttgtgtatcttgatt 60  
DB 1 TCAGATGACGAATCGCGCTACCTTGATTNGAAATCTAGGTGATGATCTTGATT 60  
QY agtttttgatactcttgctgaattcttcttagagatgcaacggtcttcaattat 120  
DB agtttttgatactcttgctgaattcttcttagagatgcaacggtcttcaattat 120  
QY agtttttgatactcttgctgaattcttcttagagatgcaacggtcttcaattat 120  
DB agtttttgatactcttgctgaattcttcttagagatgcaacggtcttcaattat 120  
QY 61 agtttttgatactcttgctgaattcttcttagagatgcaacggtcttcaattat 120  
DB 61 agtttttgatactcttgctgaattcttcttagagatgcaacggtcttcaattat 120  
QY 121 gagcccttgatgacgaagaatgacatgcatgtttgacagagtgagtaaaagta 180  
DB 121 gagcccttgatgacgaagaatgacatgcatgtttgacagagtgagtaaaagta 180  
QY 121 gagcccttgatgacgaagaatgacatgcatgtttgacagagtgagtaaaagta 180  
DB 121 gagcccttgatgacgaagaatgacatgcatgtttgacagagtgagtaaaagta 180  
QY 181 ggcgtgttgatgacgaataatgattgttgagccctatgagcttcaatgaatttga 240  
DB 181 ggcgtgttgatgacgaataatgattgttgagccctatgagcttcaatgaatttga 240  
QY 181 ggcgtgttgatgacgaataatgattgttgagccctatgagcttcaatgaatttga 240  
DB 181 ggcgtgttgatgacgaataatgattgttgagccctatgagcttcaatgaatttga 240  
QY 241 gcaagatatactgttatatgacatgcatgtagaggtatataaaagaataagacatt 300  
DB 241 gcaagatatactgttatatgacatgcatgtagaggtatataaaagaataagacatt 300  
QY 241 gcaagatatactgttatatgacatgcatgtagaggtatataaaagaataagacatt 300  
DB 241 gcaagatatactgttatatgacatgcatgtagaggtatataaaagaataagacatt 300  
QY 301 ctctcgatgacatcaattctctcttgcctaatgtaagaagtttgcacacagcgtt 360  
DB 301 ctctcgatgacatcaattctctcttgcctaatgtaagaagtttgcacacagcgtt 360  
QY 301 ctctcgatgacatcaattctctcttgcctaatgtaagaagtttgcacacagcgtt 360  
DB 301 ctctcgatgacatcaattctctcttgcctaatgtaagaagtttgcacacagcgtt 360  
QY 361 gaaactgagatgctgctcttccatcactaagttgattaccgtgaagtgctcaattc 420  
DB 361 gaaactgagatgctgctcttccatcactaagttgattaccgtgaagtgctcaattc 420  
QY 361 gaaactgagatgctgctcttccatcactaagttgattaccgtgaagtgctcaattc 420  
DB 361 gaaactgagatgctgctcttccatcactaagttgattaccgtgaagtgctcaattc 420  
QY 421 ttcgtagcccttcttccgcaatttcccaagacaagaattccttgaatgaattatgt 480  
DB 421 ttcgtagcccttcttccgcaatttcccaagacaagaattccttgaatgaattatgt 480  
QY 421 ttcgtagcccttcttccgcaatttcccaagacaagaattccttgaatgaattatgt 480  
DB 421 ttcgtagcccttcttccgcaatttcccaagacaagaattccttgaatgaattatgt 480  
QY 481 ctctgtagcgtgtgtgcatcttgaagaactcctctgtttttaaagtgaagaagaac 540  
DB 481 ctctgtagcgtgtgtgcatcttgaagaactcctctgtttttaaagtgaagaagaac 540  
QY 481 ctctgtagcgtgtgtgcatcttgaagaactcctctgtttttaaagtgaagaagaac 540  
DB 481 ctctgtagcgtgtgtgcatcttgaagaactcctctgtttttaaagtgaagaagaac 540  
QY 541 ctacacttttctgtgttcccttcttatgatacaccctcttgagaggtttagaactta 600  
DB 541 ctacacttttctgtgttcccttcttatgatacaccctcttgagaggtttagaactta 600  
QY 541 ctacacttttctgtgttcccttcttatgatacaccctcttgagaggtttagaactta 600  
DB 541 ctacacttttctgtgttcccttcttatgatacaccctcttgagaggtttagaactta 600  
QY 601 tccacctaaactatcaatgtttacagacaagctatacgtttatcccccgcgcg 660  
DB 601 tccacctaaactatcaatgtttacagacaagctatacgtttatcccccgcgcg 660  
QY 601 tccacctaaactatcaatgtttacagacaagctatacgtttatcccccgcgcg 660  
DB 601 tccacctaaactatcaatgtttacagacaagctatacgtttatcccccgcgcg 660  
QY 661 acctgtagcaaaagaagcgtgtagaactgtattatccggttttatataatacaaa 720  
DB 661 acctgtagcaaaagaagcgtgtagaactgtattatccggttttatataatacaaa 720  
QY 661 acctgtagcaaaagaagcgtgtagaactgtattatccggttttatataatacaaa 720  
DB 661 acctgtagcaaaagaagcgtgtagaactgtattatccggttttatataatacaaa 720  
QY 721 ttgatgctcatgagactaatatcttggttaattcaatgataataatacccttccct 780  
DB 721 ttgatgctcatgagactaatatcttggttaattcaatgataataatacccttccct 780  
QY 721 ttgatgctcatgagactaatatcttggttaattcaatgataataatacccttccct 780  
DB 721 ttgatgctcatgagactaatatcttggttaattcaatgataataatacccttccct 780  
QY 781 ctgtgtagcagtgagccctgtttaattagcgtactaatatcatgtatgatactgta 840  
DB 781 ctgtgtagcagtgagccctgtttaattagcgtactaatatcatgtatgatactgta 840

DB 781 ctgtgtagcagtgagccctgtttaattagcgtactaatatcatgtatgatactgta 840  
QY 841 tgccttgcaagatgacatcagatataccatttggctgtctacaataaataaac 900  
DB 841 tgccttgcaagatgacatcagatataccatttggctgtctacaataaataaac 900  
QY 901 tctattatgctatcttcttggtaagaagtaactaatgagcgttggtaagagcg 960  
DB 901 tctattatgctatcttcttggtaagaagtaactaatgagcgttggtaagagcg 960  
QY 901 tctattatgctatcttcttggtaagaagtaactaatgagcgttggtaagagcg 960  
DB 901 tctattatgctatcttcttggtaagaagtaactaatgagcgttggtaagagcg 960  
QY 961 atgcattacacaacttgcagactagcttgcagatgcccaaatgtccctgcagagctg 1020  
DB 961 atgcattacacaacttgcagactagcttgcagatgcccaaatgtccctgcagagctg 1020  
QY 961 atgcattacacaacttgcagactagcttgcagatgcccaaatgtccctgcagagctg 1020  
DB 961 atgcattacacaacttgcagactagcttgcagatgcccaaatgtccctgcagagctg 1020  
QY 1021 atccaaccccttgtaacccttgacatggttcaatgtgacatgtaacaaatgaaacgtg 1080  
DB 1021 atccaaccccttgtaacccttgacatggttcaatgtgacatgtaacaaatgaaacgtg 1080  
QY 1021 atccaaccccttgtaacccttgacatggttcaatgtgacatgtaacaaatgaaacgtg 1080  
DB 1021 atccaaccccttgtaacccttgacatggttcaatgtgacatgtaacaaatgaaacgtg 1080  
QY 1081 ttataagagtgtagtgcacttcccttataatttttttagcaagtttagcaatttact 1140  
DB 1081 ttataagagtgtagtgcacttcccttataatttttttagcaagtttagcaatttact 1140  
QY 1081 ttataagagtgtagtgcacttcccttataatttttttagcaagtttagcaatttact 1140  
DB 1081 ttataagagtgtagtgcacttcccttataatttttttagcaagtttagcaatttact 1140  
QY 1141 aatgagcagatgctccttataatatttctttaaatttcttagcgaagcagacatca 1200  
DB 1141 aatgagcagatgctccttataatatttctttaaatttcttagcgaagcagacatca 1200  
QY 1141 aatgagcagatgctccttataatatttctttaaatttcttagcgaagcagacatca 1200  
DB 1141 aatgagcagatgctccttataatatttctttaaatttcttagcgaagcagacatca 1200  
QY 1201 tcttaagatcctcactgaatttaagaacataacatttctttaaatactattagat 1260  
DB 1201 tcttaagatcctcactgaatttaagaacataacatttctttaaatactattagat 1260  
QY 1201 tcttaagatcctcactgaatttaagaacataacatttctttaaatactattagat 1260  
DB 1201 tcttaagatcctcactgaatttaagaacataacatttctttaaatactattagat 1260  
QY 1261 gtttttcgacagcagacatatccttcttgcgaatttcaagaacaccccttcccg 1320  
DB 1261 gtttttcgacagcagacatatccttcttgcgaatttcaagaacaccccttcccg 1320  
QY 1261 gtttttcgacagcagacatatccttcttgcgaatttcaagaacaccccttcccg 1320  
DB 1261 gtttttcgacagcagacatatccttcttgcgaatttcaagaacaccccttcccg 1320  
QY 1321 acgataaataataataagatactccttgaactaaattttaaatttcttcttctta 1380  
DB 1321 acgataaataataataagatactccttgaactaaattttaaatttcttcttctta 1380  
QY 1321 acgataaataataataagatactccttgaactaaattttaaatttcttcttctta 1380  
DB 1321 acgataaataataataagatactccttgaactaaattttaaatttcttcttctta 1380  
QY 1381 tgttcttgcagaagaatttctttagtcttcttgcgaatttcaagaacaccccttcccg 1440  
DB 1381 tgttcttgcagaagaatttctttagtcttcttgcgaatttcaagaacaccccttcccg 1440  
QY 1381 tgttcttgcagaagaatttctttagtcttcttgcgaatttcaagaacaccccttcccg 1440  
DB 1381 tgttcttgcagaagaatttctttagtcttcttgcgaatttcaagaacaccccttcccg 1440  
QY 1441 tggcaactcacaataattatataatccatgtaatttaagtttaaaagaagcttccca 1500  
DB 1441 tggcaactcacaataattatataatccatgtaatttaagtttaaaagaagcttccca 1500  
QY 1441 tggcaactcacaataattatataatccatgtaatttaagtttaaaagaagcttccca 1500  
DB 1441 tggcaactcacaataattatataatccatgtaatttaagtttaaaagaagcttccca 1500  
QY 1501 ttaattccaagatagaggttcaattttaaagtttggcgcacatgagtttgaanaatgt 1560  
DB 1501 ttaattccaagatagaggttcaattttaaagtttggcgcacatgagtttgaanaatgt 1560  
QY 1501 ttaattccaagatagaggttcaattttaaagtttggcgcacatgagtttgaanaatgt 1560  
DB 1501 ttaattccaagatagaggttcaattttaaagtttggcgcacatgagtttgaanaatgt 1560  
QY 1561 cagaaatttggtagtgaatttacttcccaactttagtggtcatgtagtcttg 1620  
DB 1561 cagaaatttggtagtgaatttacttcccaactttagtggtcatgtagtcttg 1620  
QY 1561 cagaaatttggtagtgaatttacttcccaactttagtggtcatgtagtcttg 1620  
DB 1561 cagaaatttggtagtgaatttacttcccaactttagtggtcatgtagtcttg 1620  
QY 1621 ggaatgcagactatcctgcaattgttcccttggcgaagttgaaataattacaact 1680  
DB 1621 ggaatgcagactatcctgcaattgttcccttggcgaagttgaaataattacaact 1680  
QY 1621 ggaatgcagactatcctgcaattgttcccttggcgaagttgaaataattacaact 1680  
DB 1621 ggaatgcagactatcctgcaattgttcccttggcgaagttgaaataattacaact 1680  
QY 1681 tgaagaacatbacttgaatgtcttagttttaaagaacaaatgcttcaabact 1740  
DB 1681 tgaagaacatbacttgaatgtcttagttttaaagaacaaatgcttcaabact 1740  
QY 1681 tgaagaacatbacttgaatgtcttagttttaaagaacaaatgcttcaabact 1740  
DB 1681 tgaagaacatbacttgaatgtcttagttttaaagaacaaatgcttcaabact 1740  
QY 1741 gttaaagtgtagaaagttgacttctcagctcagcagttgttcggaataatcatga 1800  
DB 1741 gttaaagtgtagaaagttgacttctcagctcagcagttgttcggaataatcatga 1800  
QY 1741 gttaaagtgtagaaagttgacttctcagctcagcagttgttcggaataatcatga 1800  
DB 1741 gttaaagtgtagaaagttgacttctcagctcagcagttgttcggaataatcatga 1800  
QY 1801 agcaactaaagcgtggaacttttggctatatttcaaatgttaattgttacta 1860  
DB 1801 agcaactaaagcgtggaacttttggctatatttcaaatgttaattgttacta 1860  
QY 1801 agcaactaaagcgtggaacttttggctatatttcaaatgttaattgttacta 1860  
DB 1801 agcaactaaagcgtggaacttttggctatatttcaaatgttaattgttacta 1860  
QY 1861 cttaataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1920  
DB 1861 cttaataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1920



Qy	1921	agacacatnaataaatctgpnaggttggacaaatataactctcttcttaagtlccagaa	1980
Dy	1921	agtcacatnntnaaaattgnggggttgcaaaaataaactctcttcttaaggtccaaaa	1980
Qy	1981	gagcaactatcaaccctgtctagcagcataacgtcaacgttgatcagtcacagcgtatcca	2040
Dy	1981	gagacactatcaacaccttgcgtctagccgatatacgtcacagtggtcagtcacagcgtatcca	2040
Qy	2041	gtttggagaggttctaagaacctattaccctgtctctttaaagctcgaagatgta	2100
Dy	2041	gtttggagaggttctaagaacctattaccctgtctctttaaagctcgaagatgta	2100
Qy	2101	ttaaagcgcacatctcagagtttaaatatagcacttcagcttatatgaatgagaca	2160
Dy	2101	tttaagcgcacatctcagagtttaaatatagcacttcagcttatatgaatgagaca	2160
Qy	2161	tgaagaatacatatcttaatgthcctatgcctgttccaacatgctctcaatattcgtta	2220
Dy	2161	tgaagaatacatatcttaatgthcctatgcctgttccaacatgctctcaatattcgtta	2220
Qy	2221	tcctttgcattcttaaaaaatggcactgataaaatgtggaaagtgatcctccaatcca	2280
Dy	2221	tcctttgcattcttaaaaaatggcactgataaaatgtggaaagtgatcctccaatcca	2280
Qy	2281	tttcaatctacacagaagaatcatataatcttctttaaataaagtgtggccctaagatt	2340
Dy	2281	tttcaatctacacagaagaatcatataatcttctttaaataaagtgtggccctaagatt	2340
Qy	2341	tctcagtatgtgtctattatatttccacacttcaagaactgctgtgcagaatgaatc	2400
Dy	2341	ttctcagtatgtgtctattatatttccacacttcaagaactgctgtgcagaatgaatc	2400
Qy	2401	tttgacactccacagaagaactcttaagttaaagtatctctttagactgtgtgaatgtg	2460
Dy	2401	tttgacactccacagaagaactcttaagttaaagtatctctttagactgtgtgaatgtg	2460
Qy	2461	tcatgtgtgtgacctgtccagaatttaaatcaatcccatgltcacatgcttgttgatctga	2520
Dy	2461	tcacatgtgtgtgacctgtccagaatttaaatcaatcccatgltcacatgcttgttgatctga	2520
Qy	2521	ctaacactcttaatcgaagaagaactatttgtaattaaagcctttttttgtgtctc	2580
Dy	2521	ctaacactcttaatcgaagaagaactatttgtaattaaagcctttttttgtgtctc	2580
Qy	2581	catgcttaagcgttataaaggctcctacgcctgcagactatggttcttaaatgttaagt	2640
Dy	2581	catgcttaagcgttataaaggctcctacgcctgcagactatggttcttaaatgttaagt	2640
Qy	2641	actgacaagtttaaaagttctcttctgtttacgaattagaatatataataaagccttt	2700
Dy	2641	actgacaagtttaaaagttctcttctgtttacgaattagaatatataataaagccttt	2700
Qy	2701	aaacttctcgtggaaggtgttcttaacttttataatataatagaatgactcagaact	2760
Dy	2701	aaacttctcgtggaaggtgttcttaacttttataatataatagaatgactcagaact	2760
Qy	2761	gctggaactatatactcttaagaactcagatcacagaactgttataataggttcaagt	2820
Dy	2761	gctggaactatatactcttaagaactcagatcacagaactgttataataggttcaagt	2820
Qy	2821	agtgagcgtgtatgagaacacttaagaagaacttaactatgaggtttcaacctttaact	2880
Dy	2821	agtgagcgtgtatgagaacacttaagaagaacttaactatgaggtttcaacctttaact	2880
Qy	2881	tctcttllagaatttcttcaagttatagaaaaattgtaacctatgactcgtgttcttccc	2940
Dy	2881	ttctcttllagaatttcttcaagttatagaaaaattgtaacctatgactcgtgttcttccc	2940
Qy	2941	ataaaccttcacataagtcggttctcttgacgtttccatgtaagctgttgacagagttc	3000
Dy	2941	ataaaccttcacataagtcggttctcttgacgtttccatgtaagctgttgacagagttc	3000

QY	3001	atggcggtcttcttcaaatcaatcaatgaagtcctcaacttggtagcgctgactatatt	3066
Dd	3001	ATTAGCGGTTCTTCATATATCAATAATGTCCTCACTTGAAGAGCGCTGACTATTAT	3060
QY	3061	tgcaccttgactcaacctgtaaccccaagtcactcttgattgtaactagctcactaaccca	3120
Dd	3061	TGCACCTTGACCTTAACCTTGATCCGATGTCATGTCATTGTATGTCAATGATCAACGGA	3120
QY	3121	gttcaacaatggttatcaatgctctttggaggtacaatgtaagttcacctctgccttga	3180
Dd	3121	GTTAACATGGTTATCATATCTTTTGGAGTAACAATGATGCTTCACCTGTGCTCTTGA	3180
QY	3181	tatagtttaaggcttgcaacctcccaactagcctcttgattgttttaattcaagttcaaca	3240
Dd	3181	TATAGTTTAAGGCTTGCAACCTCCACCTGACCTTTGCTGTTTATTCACAGTTCAACACA	3240
QY	3241	ccctcagcaactgttcaaccttagctctcttctgttcgcaaataglaagaattcttcgca	3300
Dd	3241	CCCTCCTGCACTGTTGACCTCTGATCTTTTGGCCGAATAGTAAGAAGTTCTTTCGCA	3300
QY	3301	taataagtgatgatacatttaagaagaatgaatcaaatatgcttatgtttgtttgac	3360
Dd	3301	TAAATAGTGATGATCATTTTAGAAATGTAATCAATATGCTGTATTTGTGTTTGATC	3360
QY	3361	tcttgaaattaaatgagttgcttgacaatttgctgttttaactgttgcataaggcttgcaca	3420
Dd	3361	TTTGGAATTAATAGTTGCTGGAACAATGTGTGTTTATGCTGTTCAGAGGCTTTCGCCA	3420
QY	3421	gggaaggcgatagtaagaagtcgggacatcccaaggccttactcttgaaggggcgggcgac	3480
Dd	3421	GGGAAGCGCATAGTAAAGTGGGCAATCCAGCGCCTTATTCCTGAAGGGCGGGCGCAC	3480
QY	3481	gttcgagcttcgtggtctcatttagaggaacatcattatataactgttaatttaaga	3540
Dd	3481	GTTGGAATTCGTGGCTGCTCATTAGAGGACATATGCTATATTAAGTATTAATTATGA	3540
QY	3541	atataaatcaactacatattttcttctglaatgttataatagaatcccaactcgtaaac	3600
Dd	3541	ATATATAACAACACTCTATATTATTTCTTGTGAATGTTAATAGGAATCCCACTGCTAAC	3600
QY	3601	ttagaanaatcacattgaataatttgaaccctaataattagttagttcaagtttaaatcca	3660
Dd	3601	TTGACAAATACCACTTGAAATATTGGAACCTTAATTAATTGTAGTGTCAAGTTTAATATCA	3660
QY	3661	aactcaatttaatttactttaaaaaaataattcctataatgaatcgtatacaggtataaatat	3720
Dd	3661	AACCTATTATTTACTTTAAAAAATAATTCTATATGATGATGTAACGATTAATAATATAT	3720
QY	3721	taaatataatgatagttgctctatatcatagctgtaagtcttaataagaatcccaaggcgctg	3780
Dd	3721	TAAATTAATATATATGTGTGCTATATATATACCTGAATAGTAAATTAATAGATCCCAAGAGCGCTG	3780
QY	3781	ctcttaactgcctaagcgctcccaaggcaagttcaactgatacttcaacttgaacaataatgggctc	3840
Dd	3781	CTCTTAACGTATATGTGTGCTATATATATACCTGAATAGTAAATTAATAGATCCCAAGAGCGCTG	3840
QY	3841	gtaataaatgttggggaatcccaatcaactgatactctcgttttgcgtgaacctctgttcaat	3900
Dd	3841	GTAATACACTGTTGTGGGATCCCTATACACTGATATCCGTGTTTGCATGCCCTCTGTTCAAT	3900
QY	3901	tgaatttcaatgatagtatctactagttttataaatcttcttataatttgaataatttaac	3960
Dd	3901	TGATTTTATATGATGTGATTAATTAAGTATTTAATAAATATCTTTATTTGCATAAATATTAC	3960
QY	3961	tggagtttaacaatgacaaggagctttacagaacaataacaataagtygacaactttccatag	4020
Dd	3961	TGGAGTTTAACAATGACAAGGAGCTTTACAGCAATATTAAGTGAGCAACCAATCTCTAGTG	4020
QY	4021	atcttgggaactctgaacaatttggtagagcttgaacctatacagatagatgcttcttgac	4080
Dd	4021	ATCTTGGGAATCTGAACAATTTGGTAGAGCTTGAACCTATATACGAATATACCTTCTGTGAC	4080
QY	4081	ctataccggaacattttaggaagaagttaacaaggtaagattctctgtatgactacaactct	4140



Db 4081 CTATACCGGACACTTGGAAAAGCTTACAAAGCTTAAGATTCTGTATGACTATCAAAATCTT 4140  
Qy 4141 cactagttctttaacttaactgaacttgatatacccttcaagtgatigtatataccacaa 4200  
Db 4141 CACTAGTTTTTAACCTATATGCAATTTGATATATCTTCAAGTGATATATATACAA 4200  
Qy 4201 ttactgataagggtctccaaacaaactgcctctctgtcccaattccaatgtaactgact 4260  
Db 4201 TTTACTGATAGGCGTCTCAACAAACAACTGCTCTGTGTCAAATCCAAATGCACTGACT 4260  
Qy 4261 aatatcaactctcaagtcctgtaagttatccgaccttccagacttgatigtgtg 4320  
Db 4261 AATATTACACTCTTCAAGTCTTAAGTATTCGACCTTCCAGATAGTTGTTGTG 4320  
Qy 4321 tggatgttcaactttaactaaataatgtaacaaaggatattcaacaacatcgact 4380  
Db 4321 TGGATGTTTCAATTTTAATTAATAATGTTCAACAGGATTTATCAAAACATCGGCTAT 4380  
Qy 4381 caggaccacgtaccggataatggtcatttctctgtttaaccattcaagtttaagtcta 4440  
Db 4381 CAGAACCAAGTACCGGATATAGGCTCATTTCTTTGTTTACACCTATACAGGTTTAATGCTA 4440  
Qy 4441 gtaatatcttaactataatggttcttactctactgagaaagctatataattttt 4500  
Db 4441 GTAATATCTTAATATATATGTTCTTACTTCTACTGCGAAAGCTATATATATTTT 4500  
Qy 4501 tctccctcaataataatcaactctgcaggtttggcaataattgaattatgtgacctg 4560  
Db 4501 TCTCTTCATATATATATATCACTTTCGAGTTTGGCAATTAATTTATGTTGAGACTG 4560  
Qy 4561 taaactggagagccctgcccctgagatctcccccatttctccaccaccctccctccac 4620  
Db 4561 TAATGAGAGCCCTGCCCCGTGATCTCCCCATTTTCTCCACCACTCCGTTCAATCCAC 4620  
Qy 4621 catcaacagtaacagcctcccaagtgatattgaatttataataatccgtaattatitta 4680  
Db 4621 CATCAACAGTACAGCCCTCAGGTGATTTAGTTTATATATATCCGTAATATATTTTA 4680  
Qy 4681 tgaactgttaaaaaatgggtgttaattcaaccagttgcgaataaagtatttctctctc 4740  
Db 4681 TGACTGTAAATATGGTGTATATTCACCAAGTTCGAAATAAAGTATTTCTCTTCTCTC 4740  
Qy 4741 ttcttatattatgaagacaanaatgtgccacgtgagactatgtctggggagtagctgc 4800  
Db 4741 TTTCTATATATATGAGAGCAAAATGATGCCACTGAGACTTCTCTGGGGAGTAGCTGC 4800  
Qy 4801 tgggtgtccttactgtttgtctgcaactgcaatgycatttgcatgtgtgagcgaagaana 4860  
Db 4801 TGGTCTGCTTACTGTTGCTGCACTGCAATGGCATTTGCAATGAGTGGCGAGAAAAA 4860  
Qy 4861 acccggaagaacatttcttgtgtgtccaggttaagtcctgttaataagatatattgaagc 4920  
Db 4861 ACCGGGAACAAATTTCTTGTATGTGCCAGGTAGTCCGTGAAATAGATATCTATTTGAGC 4920  
Qy 4921 gcttactgtctgtgagacttgtttcaactgtcatatgaattcaactcaactgagaagacc 4980  
Db 4921 GCTTACTGTCTGTGAGACTTGTTTTCACTGTCACTTAAGTAACTCACTGAGTGAAGACCC 4980  
Qy 4981 agaaagtacacttggttcaactgaagaagtttctctgcgagaattcgcaagtcgaacgga 5040  
Db 4981 AGAAGTGCACCTTGTCTCAACTGAAAGAGGTTTCTCTCGAANAATTCGAAGTCGCAACGGA 5040  
Qy 5041 taactttgataccactcttggaagaagtggtgtgtgaagttgataaggaagcccttgc 5100  
Db 5041 TACTTTTAGTACATCTTGTGGAAGAGGTGATTTGGTAAGGTATTAAGGACCCCTTGC 5100  
Qy 5101 tgaatgtcacttgaagagttaaaggcttaagaagaagaacacaaagtgagcgt 5160  
Db 5101 TGATGTGCTACCTGTAGAGAGTTAAAGGCTTAAGGAAGAAGAACACCAAGGTGCGAGCT 5160  
Qy 5161 gcaagtttcaacagaagtggaaatgattagcaatgctgtgcatcgaaatctcttgcgtct 5220  
Db 5161 GCAGTTTCAACAGAAAGTGAATGATTAAGCATGGCTGTGCATCGAAATTTCTGCGTCT 5220  
Qy 5221 acgtgttctcgtcatagcacctaccagagcggtctctgtatataccatacagtctaagtg 5280  
Db 5221 ACGTGTCTTCTGCATAGCAACCTACAGAGCGGCTCTCTGTATATCCAAACAGGCTAATGG 5280  
Qy 5281 aagtttgcgtatgtttaaagagatctcagtttaacaaattcaacttgcagaagtt 5340  
Db 5281 AAGTGTGGGTATGTTAAGAGGTATCTGATTAATCAATTTCCAACTTCCGAAAGTT 5340  
Qy 5341 tgtttgattaaataatgaataataactccctacactatgttgaagggtatataattctgag 5400  
Db 5341 TGTGTGATTAATAAATGAATATTAATCTCCCTACACTATGTAAGGAGTATATATTCGAG 5400  
Qy 5401 cagatcttattcccatgtgcaagataccagttatattgttttctcgttaattatccg 5460  
Db 5401 CAGATCTATTTCCCATTCGACAGATACAGATTAATATGTTTCTTTCGTAAATTAATCCG 5460  
Qy 5461 gttataattcttctgtatgttatttggttatatgcaaggatttcgagttcaataagtataca 5520  
Db 5461 GTATATTTCTTTCTTGTATTTGGTTATATGCAAGATTTTGAGTCTATATATATCA 5520  
Qy 5521 actggaatgataatgttattctgcaattgcaatctgtcttcaatgtgccaataatataga 5580  
Db 5521 ACTGATGCTATGTTTATTCGCAATGTAATTCCTTGTCTTCAATGAGCCAAATATATATGA 5580  
Qy 5581 tccaacttggatcaacttataataactgtgtgaagtcgaagctgttgacttcaattta 5640  
Db 5581 TTCAACTTGGAAATCATTTATATATATGTTGTAAGTACACTGTGACTTTCATCATTA 5640  
Qy 5641 attagttctcaataatcaagaatctgcctcagtagagcttaaccgaatactcaaaccttcc 5700  
Db 5641 ATTAGCTTCATTAATCAGAAATCTGCTAGTAGAGCTTACCGAATATCTTAACCTTTC 5700  
Qy 5701 ttaagccctgataataacgcgtcccaacttaattcaagttgtctgcctctcgaattc 5760  
Db 5701 TTAATGCCCCGTGTATTAATTCGTCACACTTACTTATTCAGTTTGTCTGCTCTCGAATTT 5760  
Qy 5761 ttgatctgaactgtgtgtgtgttcttcaataatgaaggtcagcatagaagcc 5820  
Db 5761 TTGATTTGACATTTGATGTCTTGTTCATCAATATATAGGCTCAGCCATCGAAGCC 5820  
Qy 5821 tccccctgtatggcaacatgaaggaagagatgtaacaaagatcttcaagggcctatcaa 5880  
Db 5821 TCCCCGTGATTTGCCACATGAGGAGAGATGTCATAGATCTTCTAGGGGCTATCTAA 5880  
Qy 5881 attgaatgcaactgtgatcccaagaattacatccgaatgtaataagctgcaataatact 5940  
Db 5881 ATTGCATGACCAATTTGATCCCAATTAATCATGTCATGTAAAGCTGCAAAATATATT 5940  
Qy 5941 attgagcgaagaatttgaagcggtgtgtgagttatttggttaagcgaagccatgagta 6000  
Db 5941 ATTGAGCAAGAAATTTGAGAGGCTGTGTGTAGGTATTTGGTTAGCTTAGGCTCATGAGATTA 6000  
Qy 6001 caaggaataccatgtttagcactgctgtlaaggggtlaacaaatgggcaatagctcccgatga 6060  
Db 6001 CAAGGATACCAATGTTAGACATGCTGTAAAGGGTACCAATGGGCAATAGCTCCGAGATA 6060  
Qy 6061 cctctcgaactggaagaagtaacagaagaacgaatgtcttggttatggaataatgctcct 6120  
Db 6061 CCTCTGACATGGAATAATCTATCAGAGAAACGATATCTTGTGTATGGAATAATGCTCTCT 6120  
Qy 6121 aagagctcattactggaacagaaggcctttgatatctgcgaccttggaacgaatgtagatgt 6180  
Db 6121 AAGGCTCATATTCTGACACAGAGGCTTTTATATTTGCTGCTGCGCTTGGCAACGATGATGATGT 6180  
Qy 6181 talgttgttgaatttgggtatgtgtccgggtgttcccttgggttaattatccacatata 6240  
Db 6181 TATGTTGTTGGAATGGGATGTGTCCGGGTGTCTCTTGGTTAATTAATTCACATATTA 6240  
Qy 6241 gtagcttaactacttgtgtgtgccccttgttttattatctcgtcgtatattgattcttagt 6300  
Db 6241 GTGCTTACTACTTGTGTGTGGCCCTTGTATTATTTCTGCTGCTGTATTTGATTTCTTAGT 6300



```

QY 6301 catgtatgcatatgacctgcttgcgaatgctttaggttaaaagccctttgaagag 6360
|||
Db 6301 CATGTTATGCAATATGACCTGCTTGGCAATGCTTTAGTTAAAGCCTTTGAAAGAG 6360
QY 6361 aaaaagttggagatgctgctgcatccctgacactgcagaacaattcatatgacagaagtt 6420
|||
Db 6361 AAAAAAGTTGGAGATGCTGCTGATCCCTGCAAGAACATTTCAATGACACGAAGTT 6420
QY 6421 gagcagcttaccagtaagtaagcttactctgtaccagaggttcgcgaatggagcgccctaa 6480
|||
Db 6421 GAGCAGCTTATTCAGATGACATTTACTCTGTACCAAGGCTTGCCAAATGGAGCGCTTAAG 6480
QY 6481 atgtcagaagtagtcgcgaatgcttgaagtgatgagccttcgagaagaagtggagagtg 6540
|||
Db 6481 ATGTGAGAGGTAGTCCGAATGCTTGAAGTGATGCGCTTCGAGAAAAGTGGAGAGACTGG 6540
QY 6541 caaaaagttgaatcatcatcaagaacgtagaattagctcacatcgaaactctgaatgg 6600
|||
Db 6541 CAAAAAGTTGAAGTCAATCCATCAAGACGTAGAAATTAGCTCCATCGAACTTCTGAATGG 6600
QY 6601 atccctagctgcagacgaatactgcatgctttgaattatctgttccaaggttaaacagca 6660
|||
Db 6601 ATCTAGACTGACAGATTAAGCTTGCACTGTTGATTAATGCTGTCACAGATTAACAGCA 6660
QY 6661 tataaatgtaataataatattttatgatt 6695
|||
Db 6661 TATAAATGTAATGAATTAATTTTATGATT 6695

```

RESULT 2

V06571  
V06571 standard; cDNA; 1814 BP.

ID V06571;  
AC V06571;  
DE 03-AUG-1998 (first entry)  
DA Daucus carota SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
RV plant breeding; ss.  
OS Daucus carota.  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 94..1755  
FT /\*tag= a  
FT /product= SERK protein

MO9743427-A1.  
PD 20-NOV-1997.  
PD 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI; 98-086529/08.  
DR P-PSDB; W47013.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 21; Pages 47-51; 123pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SO Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

```

Query Match      8.2%; Score 550.4; DB 1; Length 1814;
Best Local Similarity 80.6%; Pred. No. 9,1e-108;
Matches 722; Conservative 0; Mismatches 31; Indels 143; Gaps 1;

QY 5800 agagcgctcagcaccatcagaacccccccttgatggtccaaactagagagagatgacatag 5859
|||
Db 1044 AGAGCGCTCAGCCATCAGAACCTCCCTTGATTTGGCCAACTAGAGAGAGATGACATAGG 1103
QY 5860 atctttaggggctactcctaaattgcatgaccattgtatcccaagattatccatcgga 5919
|||
Db 1104 ATCTGTAGGGGCTTTCTTATTTGATGACATGCCATTTGATCCCAAGATTATCCATCGGA 1163
QY 5920 ttraaagctgcgaatatattatgacgaagaatttgaagctgttgaatgattttgg 5979
|||
Db 1164 TGTAAAGCTGCMAATATATTATTGACGAAAGATTTGGGCTGTGTAGGATTTGG 1223
QY 5980 gttagctaggtcattgattacagaagatcccatgttacgactgtgtaaggggtaccat 6039
|||
Db 1224 GTTAGCTAGGCTCATGAGATTACAGGATGCCATGTTTCAACTGCTTGAAGGGTACTT 1283
QY 6040 tgggcaatagctcccgagtaacctcgactggaagatcatcagaagaacgatgctt 6099
|||
Db 1284 GGGCTACATAGCTCCCGATGACTCTCGACTGGAAGTGCATCAGAGAACCGATGCTT 1343
QY 6100 tgggtatggataatgctcctctagagctactgtgagaagaggttttgatctgtcg 6159
|||
Db 1344 TGGTTATGGGATTAATGCTTTTGAAGCTCATTAAGTGCAGACAGAGCTTTTGATCTTGCTG 1403
QY 6160 ccttggaacgatgatgatgattatgttggattggattggtatgtygccgggtgtccctt 6219
|||
Db 1404 CCTTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
QY 6220 ggttaatttcaacatattagtgcttactactgtgttgcccttgctttattcc 6279
|||
Db 1440 ----- 1440
QY 6280 tgcctgattgattctcttagtcatgttaagcatattgacctgttgcaatgctttag 6339
|||
Db 1440 -----G 1440
QY 6340 gttaaagccttttgaagaagaagaattggaatgctgtgctgcactgcagactgcagaag 6399
|||
Db 1441 GTTAAAGCCTTTTGAAGAAAGAAAAGTTGGAAGTCTGCTGATCCTTGACCTGGAAGAAC 1500
QY 6400 aattacattgacacagaagtgtagcagcttattcagaatgacattactctgtaccagaagt 6459
|||
Db 1501 AATTACATTGACACAGAAAGTTGAGCGCTTATTCAAGTATGATCTGTATGCCAGGGT 1560
QY 6460 tggccaatggagggccttaagatgtcagaagtgatccgaaatgcttgaaggtatggcctt 6519
|||
Db 1561 TCGCCAATGGAGGGGCTTAAAGTGCAGAGTGCAGATCTTGAAGGTATGGCCTT 1620
QY 6520 ggcgaagaagtggagcagatgggcaaaaagtgaatcatccatcagaagcgttagaattagct 6579
|||
Db 1621 GCAGAAAAGTGGGACAGAGTGGCAAAAAGTGAAGTATCATCAACAGACTTAATTAAGCT 1680
QY 6580 ccacatcgaaacttctgaatgatactagactgcagacagaatactgcatgcttttaatta 6639
|||
Db 1681 CCACATCGAAGCTCTGAAGAGATCCTAGACTCGACAGATTAAGTTCATTAATTAATTA 1740
QY 6640 tctgttccaagaataacagcatataaatatgtaataattattttatgatt 6695
|||
Db 1741 TCTGTCTCAAGATACAGCATATTAATATGATGAATTAATTTTATGATT 1796

```

RESULT 3

V06585  
V06585 standard; DNA; 4081 BP.  
AC V06585;  
DE 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK gene.







OY	4984	agtcacacttgcgtcaacttgaagaggtttctctgcgcgaagattgcgaatcgcgaacggttac	5043
Db	3321	AGTTCATCTCGGACAGCTCAAGAGGCTTTCTTTGCGGAGCTACAAAGTGGCAGTATGC	3380
OY	5044	tttagtacca-----tccttggaaagagtgagatttggtaagggtgtataaggagacgct	5097
Db	3381	GTTTAGTAACAGAACATTTTGGCGAGAGGTGGGTTTGGGAAGTCTCAAGGGGACGCTT	3440
OY	5098	tgctgatggtctcactgttagcagtttaagggtcttaagagaagaacacacagtgcgga	5157
Db	3441	GGCAGACGGAACCTTCTTCTCTCAAGAGACTGAAAGGAGAGCCAACTCCAGGGGAGGA	3500
OY	5158	gctcagattccaacagaagtggaaatgatattagcatgctgtgcacgcgaactctctcgcg	5217
Db	3501	GCTCCAGTTTCAAAACAGAAATAGATGATGATGAGTATGCGAGTTCATCGAAACCTGTTGAG	3560
OY	5218	tctacgtggtttctgtcatgacaccctccgcagcgctctctgtatatccatacatgctaa	5277
Db	3561	ATTACGAGGGTTTCTGTTGTGACCGACCGAGAAATGCTTGTTGATCCTTACATGGCCAA	3620
OY	5278	tggaagtgttgcgcgatgttaagagatctcagttacaattacatacctaacttgcagaa	5337
Db	3621	TGGAAAGTCTGCTTCTCGTCTCAGAGTAAATAAACTAAACATTAACATTTGTGCTCTC	3680
OY	5338	gtttgtttgattaaataaataataactccctacacatctgttaagggtgtataattctt	5397
Db	3681	TCTCAATTACTTGTGACGTGAAGTGTTTTTCATGTTTCTTATATGAGTTCATATTGTT	3740
OY	5398	g 5398	
Db	3741	G 3741	
RESULT	4		
FT	V06591		
ID	V06591	standard: cDNA to mRNA; 2089 bp.	
AC	V06591:		
DT	03-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK gene.		
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;		
KW	plant breeding; ds.		
OS	Arabidopsis thaliana.		
FH	Key	location/Qualifiers	
FT	CDS	195..2072	
FT		/*tag= a	
PN	W09743427-A1.	/product= SERK protein	
PD	20-NOV-1997.		
PE	13-MAY-1997: E02443.		
PR	14-MAY-1996: GB-010044.		
PA	(NOVS ) NOVARTIS AG.		
PI	De Vries SC, Hecht VFG, Schmidt EDI, Van Holst GJ;		
DR	WPI; 98-086529/08.		
DR	P-ESDB: W47023.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 27: Pages 91-95; 123pp: English.		
CC	The sequence is that encoding SERK, a putative receptor kinase. It may be used as part of a method of producing apomictic seeds comprising: (a) transforming plant material with a nucleotide sequence encoding a protein which in active form in a cell or cell membrane renders the cell embryogenic; (b) regenerating the transformed material into plants or carpel-containing plant parts; and (c) expressing the sequence in the vicinity of the embryo sac. The apomictic seeds and embryos thus produced can be developed into plant progeny. This is useful in plant breeding programs. Controllable and reproducible apomixis provides many advantages in plant improvement and cultivar development in the case that sexual plants are available as crosses with the apomictic plant. Apomixis provides for true-breeding, seed propagated hybrids and could shorten and simplify the breeding process so that selfing and progeny testing to produce and/or stabilise a desirable gene combination could be eliminated.		
CC	Apomixis allows plant breeders to develop cultivars with		

CC	Specific stable traits for such characteristics as height,	
CC	seed and forage quality and maturity	
SQ	Sequence 2085 BP; 568 A; 419 C; 503 G; 599 T;	
	Query Match 3.7%; Score 248.8; DB 1; Length 2089;	
	Best Local Similarity 60.3%; Pred. No. 8e-44;	
	Matches 523; Conservative 0; Mismatches 202; Indels 143; Gaps 1	
QY	5800 agagagctcagccatcagcaaacctcccctcgtatgtggccaactgagagagatgtcactagg 5859	
DB	1361 AGAAGGCGCACCGTCAACAACCTCCGGTTGATTTGGCCAACGGGGAAGAACTGGCGTTAGG 1420	
QY	5860 atctctcaggagccctatcctaattgatcagcatctgtatcccaagattatccatcgcgga 5919	
DB	1421 CTCAGCTCGAGTTTGTCTTACCTACATATATACCTGCATCCGAAAGATCATTCACCGGTGA 1480	
QY	5920 tgtaaaagctgcaatatattatttggacgaaagatttgaagcctgtgtagtgtaatttgg 5979	
DB	1481 CGTAAAGCGAGCAAAACATCCTCTTTAGACGAAMAATTCGGACGGGTGTTGGAGATTTCGG 1540	
QY	5980 gttgacgagctatgatattacaaggatcccatgttgcattgttcgactctgtgaaggggtaccat 6039	
DB	1541 GTTGCGAAGAGCTTTATGGACATTAATAACACTACCTGACAAACAGACAGCTCGTGACACAT 1600	
QY	6040 tgggacataagctcccgagfacctcccgactcgagttgaagatccatcaagaaagacgcatgtcct 6099	
DB	1601 CGGTACATCGCTCCAGAAATATCTTCAACCGGAAATCTTGAGAAACCGACGTTT 1660	
QY	6100 tggttatggagataatgtcctctagaagctatctcctggacagagaggtcttattgtctcg 6159	
DB	1661 CGGATAGCGGATCTATGCTTCTTAGAACATTAATCAAGAGACAAGAGCTTTGCAATCCGCTCG 1720	
QY	6160 ccttgagcaagatgatgatgttattgtgttgattgggtgatgtgtgcccgggtgtccttt 6219	
DB	1721 GCTAGCTACAGACACACACACTCATGTTACTTTGACG----- 1757	
QY	6220 ggttaattalttcacatattagtgcttactacttgtgtgccccttgttttatttcc 6279	
DB	1757 ----- 1757	
QY	6280 tgcctgatttgaattccttagtcaatgtatgacatattgacgcgtcttgaatgtcttttag 6339	
DB	1757 -----g 1757	
QY	6340 gttaaaagccttttgaagaagaaaggttggagatgtgtgtcattccttgaagcttgaagac 6399	
DB	1758 GTGAAGAGATTTGTTGAAGSAGAAAGAAAGCTAAGAGTGTATAGTGGATCCAGATCTTCAACA 1817	
QY	6400 aattacattgcaagaagaagtttggacagcttatccaagtatgacatctctgtaccacaggt 6459	
DB	1818 AACTATCGAGGAGAGAAAGACTGGAAACAAGTAAACAAGTGGCGCTTCTATGGACCAAGGA 1877	
QY	6460 tcggcaaatggagcgccctaagatgctcaagaggtatgtccgaatgcttgaaggtgcatggcctt 6519	
DB	1878 TCACCAATGGAAACACCAAAAGATGTCTGAAGTTGTATAGSAGTCTTGAAAGAGATGGGCTT 1937	
QY	6520 gcaagaagaatgggacgaggtggcaaaaagttgaagtcataccatcaaagacgtlaaattagct 6579	
DB	1938 GCGGAGAAATGGGACGATAGCAAAAAAGTTGAGATTTTATAGGAAGAATGATGATTGTAGT 1997	
QY	6580 caacatgcaacttgaattgatctccatgagctctgacagatcaacttgcattcgtttgaatta 6639	
DB	1998 CCTAATCTTAACCTGTGATGGATTTCTTGATTCTACTTACAAATTTGCACGCCGTTGAGTTA 2057	
QY	6640 tctggtccaagataaacaagacataaaaa 6667	
DB	2058 TCTGGTCCAAAGCTAAAAAATAAAAAA 2085	
RESULT	5	
ID	X23533 standard; DNA; 3842 BP.	
	X23533	



AC X23533;  
 DT 17-JUN-1999 (first entry)  
 DE Tomato Xa21 clone TRK2 DNA fragment.  
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice; TRK2;  
 KM plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
 OS Lycopersicon esculentum.  
 PN WO9909151-A2.  
 PD 25-FEB-1999.  
 PF 17-JUL-1998; U14841.  
 PR 13-AUG-1997; US-910386.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 DR WPI; 99-204431/17.  
 PT New RRR polynucleotides and nucleic acid constructs - used for  
 PS generating transgenic plants resistant to Xanthomonas  
 CC Claim 14; Page 61-62; 67pp; English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the Oryza longistaminata and Oryza sativa receptor  
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
 CC rice or tomato.  
 SQ Sequence 3842 BP; 984 A; 719 C; 894 G; 1244 T;

Query Match 1.7%; Score 115; DB 1; Length 3842;  
 Best Local Similarity 56.5%; Pred. No. 2.2e-15;  
 Matches 214; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 5828 gattggccaactaaggagagatgtacatgacatctcttggggcctacttaattgcat 5887  
 DB 2975 GACGTGGAGCGTTCTTCAACAAATGCTTGGATGACCCCGTGCATCTGCTTACCTGCAT 3034  
 QY 5888 gaccattggtaccacagattatccatgcgcagttaaaagctgaataatattatgagac 5947  
 DB 3035 GATCAGTGTGACACACGCTGCTTCATCGTATGTGAAGCCGACACATTTATTTGAT 3094  
 QY 5948 gaagaatttgaagctgtttagtatttgggttagtgcctagtcacatgaagat 6007  
 DB 3095 GAGCAGTATATGATATTTATCTGATTTGGTTGGCTGATGACGGAACCTTCAGAG 3154  
 QY 6008 accatgttgaactgtcgttaagggtgacattgagcaataagctccggatccctctg 6067  
 DB 3155 ACCATGCACTACTGCTGCTGCGGACAACTTTGATATGCTCTCTGATATGCCATG 3214  
 QY 6068 actggaagatcatcagaagaacacagatcttggttatagggaataatgctcctagaagctc 6127  
 DB 3215 ACTTGCCGCGTCTGCGACAGCTGATGTATACAGTTATGGGGTTGTGTTGATGATTA 3274  
 QY 6128 attactgacagagagctttagatctgctcgccttgcgaagatgtatgtatgtt 6187  
 DB 3275 ATATCAGTAAAGAAAGCACTGATCCGCTTCTTCTTATGGAATGATTCATATATT 3334  
 QY 6188 ttgagttggtatgttctc 6206  
 DB 3335 GTAGCTTGGCATGCATGC 3353

RESULT 6  
 T31307  
 ID T31307 standard; cDNA; 1554 BP.  
 AC T31307;  
 DT 15-NOV-1996 (first entry)  
 DE Tomato RRR gene clone TRK1.  
 KW Xa21; RRR; disease resistance; Xanthomonas; blight; rice;  
 KM tomato receptor kinase 1; TRK1; transgenic plant; crop protection;  
 KM ss.  
 OS Lycopersicon esculentum.  
 PN WO9622375-A2.  
 PD 25-JUL-1996.

PF 17-JAN-1996; U00717.  
 PR 17-JAN-1995; US-373374.  
 PR 07-JUN-1995; US-475891.  
 PR 29-SEP-1995; US-004645.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Ronald PC, Song W, Szabo V, Wang G;  
 PI WPI; 96-354532/35.  
 DR P-PSDB; W03186.  
 PT Plant disease resistance gene Xa21, codes for RRR polypeptide -  
 PS useful for conferring resistance to Xanthomonas in rice and tomatoes  
 CC Example 4; Page 33; 47pp; English.  
 CC cDNA clones TRK1 (T31307) and TRK1 (T31308) were isolated from  
 CC a tomato cDNA library by screening with probes generated by PCR  
 CC amplification of tomato cDNA using primers (T31301-06) based on the  
 CC rice disease resistance Xa21 gene (T31300). They belong to the RRR  
 CC family of disease resistance genes. TRK1 (tomato receptor kinase 1)  
 CC is present as 1 or 2 copies in the tomato genome and 1 copy maps to  
 CC the short arm of chromosome 1 in the proximity of a resistance gene  
 CC to Xanthomonas campestris pv. vesicatoria Rxi. It encodes an RRR  
 CC protein (W03186) that is an important component of plant signal  
 CC transduction pathways leading to a defence response. The gene is  
 CC useful for engineering disease resistance in tomato and other plants.  
 SQ Sequence 1554 BP; 477 A; 303 C; 348 G; 426 T;

Query Match 1.6%; Score 109.6; DB 1; Length 1554;  
 Best Local Similarity 60.8%; Pred. No. 2.3e-14;  
 Matches 197; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 5829 attggccaactaaggagagatgtcactagatctctttagggcctacttaattgcat 5888  
 DB 986 AATGGGAACCTAGAGCTAAATATCCATAGAGCTGCAGAGGCGCTTCTTATTTGACAC 1045  
 QY 5889 accattgtatcccaagattatccatgcgcagttaaaagctgaataatattatgagac 5948  
 DB 1046 ACGATTGCTCCCTATGATTAATCCACCGCATGTGCAAGTCCAAATATTTGTGAAT 1105  
 QY 5949 aagaatttgaagctgtttagtatttgggttagtgcctagtcacatgaagat 6008  
 DB 1106 CTGAACCTTGAAAGCTATTTGATATTTGATATGACCAAGTACTTTCGTAACAAATGCTA 1165  
 QY 6009 cccatg---ttagactgctgtaagggttacattgagcaataagctccggatccctc 6065  
 DB 1166 CCTGTAGTGCATGCTCAATTCAGAGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225  
 QY 6066 cgactggaagatcatcagaagaacacagatcttggttatagggaataatgctcctagaagc 6125  
 DB 1226 ACACGCTGAAATGATGAGAAAGCGATGTATAGCTTTGGAGTGTGTTGGAGC 1285  
 QY 6126 tcatcactgacagagagcttct 6149  
 DB 1286 TTTATACAGAGACGAGGCGCAGTAG 1309

RESULT 7  
 X23532  
 ID X23532 standard; DNA; 3293 BP.  
 AC X23532;  
 DT 17-JUN-1999 (first entry)  
 DE Tomato Xa21 clone TRK1 DNA fragment.  
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice; TRK1;  
 KM plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
 OS Lycopersicon esculentum.  
 PN WO9909151-A2.  
 PD 25-FEB-1999.  
 PF 17-JUL-1998; U14841.  
 PR 13-AUG-1997; US-910386.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 DR WPI; 99-204431/17.  
 DR P-PSDB; W93600.  
 PT New RRR polynucleotides and nucleic acid constructs - used for



PT generating transgenic plants resistant to Xanthomonas  
 PS Claim 14; Page 59-60; 67pp; English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the Oryza longistaminata and Oryza sativa receptor  
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
 CC rice or tomato.  
 CC Sequence 3293 BP; 937 A; 634 C; 701 G; 994 T;

Query Match 1.6%; Score 109.6; DB 1; Length 3293;  
 Best Local Similarity 60.8%; Pred. No. 3e-14;  
 Matches 197; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

OY 5829 attggccaattgagagagattgacatgattcttcttaaggcctataaattgcatg 5888  
 DB 2406 AATGGGAACAGTGGCTAAATAATAGCAATAGAGCTGCGCAAGGCGCTTCTATTGTGCACC 2465  
 OY 5889 accattgtagtcccaagattatccatgcgatgtaaaagctgcaatatatttgagc 5948  
 DB 2466 AGATTGTCTCCCTTGTGTAATCCACCGCATGTCACATCCACAAATATATGTGTGACT 2525  
 OY 5949 aagaattgaagctgttagtgatttggttgtagctagagctcatgattacaagata 6008  
 DB 2526 CTGAACTTGAACCTATGTTGAGATTTTGATTAAGCAATACTTTCGTATACAAATGCTA 2585  
 OY 6009 cccattg---taagactcctgtaagggtacacattgggcacatgctccgagtagctc 6065  
 DB 2586 CCTGTGAGTGCATGCTGCAATTCGAGATCTATAGCTATGCTTCACGAATATGCAT 2645  
 OY 6066 cgactcggaagatcatcaagagagacgatgtcttggttagtgagataagctccctag 6125  
 DB 2646 AACCGCTGAATTTGATGAGAAACCGATGTGTATAGCTTTGAGTGTGTGAGAC 2705  
 OY 6126 tcattactggacagagggcctttg 6149  
 DB 2706 TTATAACAGCAGCAAGGCCAGTAG 2729

## RESULT 8

RESULT 8  
 ID T62124 standard; cDNA to mRNA; 3176 BP.  
 AC T62124;  
 DT 10-JUN-1997 (first entry)  
 DE Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.  
 KW Plant; morphogenesis; regulation; short; stem; alteration;  
 KW inflorescence; extraneous; gene; expression; transformation;  
 KW increase; control; form; length; ds.  
 OS Arabidopsis thaliana.  
 FH Key  
 FT Location/Qualifiers  
 FT cds 51..2981  
 FT /tag= a  
 FT /note= "plant morphogenesis regulatory protein"  
 PN J09056382-A.  
 PD 04-MAR-1997.  
 PE 24-AUG-1995; 216187.  
 PR 24-AUG-1995; JP-216187.  
 PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.  
 PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.  
 DR WPI: 97-206629/19.  
 DR P-PSDB; W13408.  
 PT DNA encoding plant morphogenesis regulatory protein - useful to  
 PT yield plants with short stems or altered inflorescence  
 PS Claim 1; Pages 6-10; 17pp; Japanese.  
 CC The present sequence encodes an Arabidopsis thaliana plant  
 CC morphogenesis regulatory protein (MRP), which can be used to yield  
 CC a plant with, e.g. short stems or altered inflorescence. The MRP  
 CC acts on a plant at a specific site for a specific period, and can  
 CC therefore be used to regulate extraneous gene expression in a  
 CC plant. The MRP's cDNA or genomic DNA can be used to transform a  
 CC plant to increase its MRP expression, and therefore control the

CC form (particularly stem length) of the plant.  
 SQ Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match 1.6%; Score 108.2; DB 1; Length 3176;  
 Best Local Similarity 58.1%; Pred. No. 5.8e-14;  
 Matches 191; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

OY 5823 cccctgattgccaactagagagagattgcactagatcttcttagggcctataat 5882  
 DB 2272 CTCTGATTGGGACACACGCGCTTATGATATGATGTCAGACAAAGCTTATGATTC 2331  
 OY 5883 tgcattgacctgtagtcccaagattatccatgcgatgtaaaagctgcaatatatt 5942  
 DB 2332 TACACCAATGACTGTATGAGGCTGCTTGAACAGATTTTGGAAATAGCAAAAGCTTGTGTCT 2391  
 OY 5943 tggacgaagaatttgaagctgttgtagtgatttgggttagctagagctcatgattaca 6002  
 DB 2392 TGGACAAAGACTTTAGAGGCTGCTTGAACAGATTTTGGAAATAGCAAAAGCTTGTGTCT 2451  
 OY 6003 aggatgccatgtagtgcagctgtagaggggtacacattgggcacatagctccgagtagc 6062  
 DB 2452 CAAAGTCACATACCTTCAACTTACGATGATGGGACAGATAGGTATCAATGAGACCCGAGTAG 2511  
 OY 6063 tctgcactggaagatcatcaagagagacgatgtcttggttagtgagataagctccctag 6122  
 DB 2512 CTGCACTTTCACGGCTCAGTGAAGAAATCCGATGCTATGAGAAATAGTCTTCTTG 2571  
 OY 6123 agctcattactggaacagagggcctttgat 6151  
 DB 2572 AGCTGTATACCCGAAAGAACCCGTTGAT 2600

## RESULT 9

RESULT 9  
 ID T62125 standard; DNA; 9295 BP.  
 AC T62125;  
 DT 10-JUN-1997 (first entry)  
 DE Arabidopsis thaliana plant morphogenesis regulatory protein DNA.  
 KW Plant; morphogenesis; regulation; short; stem; alteration;  
 KW inflorescence; extraneous; gene; expression; transformation;  
 KW increase; control; form; length; ds.  
 OS Arabidopsis thaliana.  
 FH Key  
 FT Location/Qualifiers  
 FT exon 1803..1881  
 FT /tag= a  
 FT 1882..2227  
 FT /tag= b  
 FT 2228..2366  
 FT /tag= c  
 FT 2367..2467  
 FT /tag= d  
 FT 2540..2643  
 FT /tag= e  
 FT 2468..2539  
 FT /tag= f  
 FT 2644..2715  
 FT /tag= g  
 FT 2716..2809  
 FT /tag= h  
 FT 2810..2878  
 FT /tag= i  
 FT 2879..2968  
 FT /tag= j  
 FT 2969..3040  
 FT /tag= k  
 FT 3041..3118  
 FT /tag= l  
 FT 3119..3190  
 FT /tag= m  
 FT 3191..3266  
 FT /tag= n  
 FT 3267..3338  
 FT exon



```

FT      Intron      /tag- O
FT      3339. .3421
FT      /tag- P
FT      3422. .3493
FT      /tag- Q
FT      3494. .3586
FT      /tag- R
FT      3587. .3655
FT      /tag- S
FT      3656. .3740
FT      /tag- T
FT      3741. .3812
FT      /tag- U
FT      3813. .3888
FT      /tag- V
FT      3889. .3960
FT      /tag- W
FT      3961. .4048
FT      /tag- X
FT      4049. .4120
FT      /tag- Y
FT      4121. .4209
FT      /tag- Z
FT      4210. .4281
FT      /tag- aa
FT      4282. .4349
FT      /tag- ab
FT      4350. .4421
FT      /tag- ac
FT      4422. .4508
FT      /tag- ad
FT      4509. .4580
FT      /tag- ae
FT      4581. .4706
FT      /tag- af
FT      4707. .4778
FT      /tag- ag
FT      4779. .4860
FT      /tag- ah
FT      4861. .4932
FT      /tag- ai
FT      4933. .5018
FT      /tag- aj
FT      5019. .5090
FT      /tag- ak
FT      5091. .5176
FT      /tag- al
FT      5177. .5248
FT      /tag- am
FT      5249. .5412
FT      /tag- an
FT      5413. .5481
FT      /tag- ao
FT      5482. .5576
FT      /tag- ap
FT      5577. .5648
FT      /tag- aq
FT      5649. .5726
FT      /tag- ar
FT      5727. .5800
FT      /tag- as
FT      5801. .5882
FT      /tag- at
FT      5883. .6011
FT      /tag- au
FT      6096. .6443
FT      /tag- av
FT      6012. .6095
FT      /tag- aw
FT      6444. .6519
FT      /tag- ax
FT      6520. .5412
FT      /tag- ay

FT      Intron
FT      6891. .6974
FT      /tag- az
FT      6975. .7328
FT      /tag- ba

J09056382-A.
PD 04-MAR-1997.
PF 24-AUG-1995; 216187.
PR 24-AUG-1995; JP-216187.
PA (MITS-) MITSUI GROSU SHOKUBUTSU BIO KENKUSHO.
PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
DR WPI; 97-206629/19.
PT DNA encoding plant morphogenesis regulatory protein - useful to
PT yield plants with short stems or altered inflorescence
PS Claim 6; Pages 12-15; 17pp; Japanese.
CC The present sequence encodes an Arabidopsis thaliana plant
CC morphogenesis regulatory protein (MRP), which can be used to yield
CC a plant with, e.g. short stems or altered inflorescence. The MRP
CC acts on a plant at a specific site for a specific period, and can
CC therefore be used to regulate extraneous gene expression in a
CC plant. The MRP's cDNA or genomic DNA can be used to transform a
CC plant to increase its MRP expression, and therefore control the
CC form (particularly stem length) of the plant.
SQ Sequence 9295 bp; 2873 A; 1581 C; 1602 G; 3239 T;

Query Match 1.6%; Score 107.2; DB 1; Length 9295;
Best Local Similarity 55.7%; Pred. No. 1.3e-13;
Matches 205; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 5823 cccctgattgccaactaggagagagattgcactagatctcttagggcctatctaatt 5882
DB 6538 ctccttgattgggacacacgaccttaagatgcatatgctgcagcacacagctttacttattc 6597
QY 5883 tgcattgaccattgtatcccaagattatccatgcgcatgtaaaagctgcaatatattat 5942
DB 6598 tACACCATGACTGTAGTCCAGAGATCATTCACAGAGAGTGAAGTCCGCAACATTCTCT 6657
QY 5943 tggacgaagaattttaggctgttgtagtgatttgggttagagcttaggtcattgata 6002
DB 6658 TGGACAAGAGCTTAGAGGCTCGTTTGACAGATTGTGGAATGCGAAGAGCTTGCTGTGT 6717
QY 6003 aggatcccatgttaacgactgctgtaagggtaccattgggacactagctccgagatcc 6062
DB 6718 CAAAGTCACATCTTCACTTCACTGATGAGGCGACATGAGTACTCAATCCCGAGATAG 6777
QY 6063 tctcgactggaagatcatcagaagaccgagctcttgattgaggaataatgctcctag 6122
DB 6778 CTCGCACCTTCACGCGCTCCTACTGAGAAATCCGATGCTACAGTATGGAATATGCTTCTG 6837
QY 6123 agctcaatactggacagagagcttttgatcttgcgccttgcgacgagatgatgtta 6182
DB 6838 AGTTGTTAACCGGAGAGAAAGCGGTGATGAGCAATCCATCTCCACCATCTGTTGTT 6897
QY 6183 tgtgtgtg 6190
DB 6898 CTTTCTTG 6905

RESULT 10
X07356
ID X07356 standard; DNA; 4104 BP.
AC X07356;
DT 21-MAY-1999 (first entry)
DE Arabidopsis steroid receptor Bin1 DNA.
KW BIN1; steroid receptor; receptor kinase; transgenic plant;
KW brassinosteroid; disease resistance; crop protection;
KW contraceptive; ss
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 97..3687
FT PN W09859039-A1.
FT 30-DEC-1998.

```



PF 24-JUN-1998; U13100.  
 PR 24-JUN-1997; US-881706.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PI Chory J, Li J;  
 DR WPI: 99-081275/07.  
 P-PSDB: M97819.  
 PT New receptor kinase Bin1 involved in brassinolide signalling -  
 PT useful for promoting increased yield and disease resistance in  
 PT plants and for modulating oocyte maturation  
 PS Claim 8; Page 49-52; 72pp; English.  
 CC This DNA sequence codes for a novel plant steroid receptor kinase,  
 CC designated Bin1 (see M97819), which is involved in the pathway for  
 CC the synthesis of the plant steroid hormone, brassinolide. 18 new  
 CC Arabidopsis dwarf mutants were identified that lacked the ability  
 CC to respond to brassinolide, and were named bin mutants. The bin1  
 CC mutations were used to map the gene to a small interval on  
 CC Arabidopsis chromosome 4. Bin1 was cloned by map-based cloning.  
 CC The Bin1 polynucleotide was identified within this interval by  
 CC sequencing the wild-type and mutant alleles of this nucleic acid.  
 CC Overexpression of Bin1 in transgenic plants provides plants  
 CC characterized as having enhanced disease resistance, increased  
 CC plant yield or vegetative biomass and increased seed yield.  
 CC Expression of Bin1 may also increase resistance to pesticides.  
 CC Inhibition of Bin1, e.g. using antisense oligonucleotides (AON), is  
 CC used to render plants male-sterile, and to reduce their stature or  
 CC yield, e.g. for creating dwarf varieties. Since Bin1 homologues  
 CC may be involved in regulation of the menstrual cycle and uterine  
 CC function, Bin1 antibodies and AON may be useful as contraceptives,  
 CC for improving success of in vitro fertilisation and to prevent  
 CC premature labour. Transgenic animals are also provided, and are  
 CC models for studying steroid-receptor interactions or can be used  
 CC to screen for therapeutic agents.  
 SQ Sequence 4104 BP; 1061 A; 870 C; 930 G; 1243 T;

Query Match 1.3%; Score 86.2; DB 1; Length 4104;  
 Best Local Similarity 53.9%; Pred. No. 2.9e-09;  
 Matches 200; Conservative 0; Mismatches 168; Indels 3; Gaps 1;

OY 5825 cctatgagccaactgagagagagatgacatgacatcttcaggggcctactaattg 5884  
 DB 3028 CTTAAATGTGTCACACGCGGAGATTGCCATGATCAGTACGAGGCTTTCCTT 3087  
 OY 5885 catgacattgtagtcccaagattccatcgcatgtaaaagctgcaaatattattg 5944  
 DB 3088 CACCAACACGCAAGTCGCGATATCATCCACAGACATGAAATCCAGTATGTGCTT 3147  
 OY 5945 gacgaagaattgagagctgtgtagtattggttgtagttaggtcattgataaag 6004  
 DB 3148 GATAGAAATTTGGAAGCTCGGTTTCAGATTTCGATGCGACGCTGATGCGATG 3207  
 OY 6005 gataccatgttca---cgactgctgtaagggtacacattgggacacatagctccagatg 6061  
 DB 3208 GATACGATTTAAGCGTCATCATATTAGCTGATACACGGGTTACGTCCTCCAGAGTAT 3267  
 OY 6062 ctctcgactggaagatcagagagacgcatgcttctgttatggataatgctccta 6121  
 DB 3268 TACCAAGTTTCAGAGTTTACACAAAGAGACGTTTATATGTTACGTTGCTTACTC 3327  
 OY 6122 gagctcattactgagacagagagctttgatactctgccttcgcttcgacagatgatgtt 6181  
 DB 3328 GAGCTACTTCACGGGTAAACGGCAACGATTCACCGATTTCGATATTAACAACCTGTGT 3387  
 OY 6182 atgtgtgtga 6192  
 DB 3388 GGATGGGTGA 3398

RESULT 11  
 X23531  
 ID X23531 standard; cDNA: 3045 BP.  
 AC X23531;  
 DT 17-JUN-1999 (first entry)

DE Maize Xa21 gene DT4 cDNA fragment.  
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice; DT4;  
 KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
 OS Zea mays.  
 PN WO909151-A2.  
 PD 25-FEB-1999.  
 PF 17-JUL-1998; U14841.  
 PR 13-AUG-1997; US-910386.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 DR WPI: 99-204431/17.  
 PT New RRR polynucleotides and nucleic acid constructs - used for  
 PT generating transgenic plants resistant to Xanthomonas  
 PS Claim 12; Page 58-59; 67pp; English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the Oryza longistaminata and Oryza sativa receptor  
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
 CC rice or tomato.  
 SQ Sequence 3045 BP; 717 A; 739 C; 744 G; 735 T;

Query Match 1.2%; Score 80.4; DB 1; Length 3045;  
 Best Local Similarity 53.7%; Pred. No. 4.5e-08;  
 Matches 190; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

OY 5825 cctgattggccaactgagagagatgacatgacatcttcaggggcctactaattg 5884  
 DB 2254 CTTGATTGGGCAACAAGAAATTGCAATTGCGCTCGCAAGAGAGCTGCTTCTC 2313  
 OY 5885 catgacattgtagtcccaagattccatcgcatgtaaaagctgcaaatattattg 5944  
 DB 2314 CACCAATGTTGCACTCCACACATACACGCGGACATGAAATGCAACGATGCTT 2373  
 OY 5945 gacgaagaattgagagctgtttagtattggttgtagttaggtcattgataaag 6004  
 DB 2374 GAGCATTAATCTCATGCTCAGTACGATTCGATTCGGAATGGCGGCTCGTAAATGCTGT 2433  
 OY 6005 gataccatgttca---gactgctgtaagggtacacattgggacacatagctccagatg 6061  
 DB 2434 GATCTCATCTTAACCGTGGAGCAAGCTTTAGGAACACCTGTTATGTCGCCAGATGAC 2493  
 OY 6062 ctctcgactggaagatcagagagacgcatgcttctgttatggataatgctccta 6121  
 DB 2494 TTCCAGTCGTTTATTTGACACACTAAGGCGACGCTACAGCTAAGGCGTTGTTCTCTG 2553  
 OY 6122 gagctcattactgagacagagagctttgatactctgccttcgcttcgacagatgat 6175  
 DB 2554 GAGCTTCTCTCAAGGAAAAAACCAATCAATCCGACTGATTCGCGACAAATATAT 2607

RESULT 12  
 V06587  
 ID V06587 standard; cDNA to mRNA: 981 BP.  
 AC V06587;  
 DT 03-AUG-1998 (first entry)  
 DE Arabidopsis thaliana SERK LRR homologous EST clone.  
 KW receptor kinase; apomictic; seeds; production; embryos;  
 KW plant breeding; laucine-rich repeat; ss.  
 OS Arabidopsis thaliana.  
 FH Key  
 FT CDS Location/Qualifiers  
 FT 104..760  
 FT /tag= a  
 FT /note= "Shows high homology to SERK"  
 PN WO9743427-A1.  
 PD 20-NOV-1997.  
 PF 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VRG, Schmidt EDL, Van Holst GJ;



DR WPI: 98-086529/08.  
 DR P-PSDB: W47019.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 28; Pages 75-77; 123pp; English.  
 CC The sequence is that of an EST clone showing high homology to  
 CC SERK LRR (leucine-rich repeat) sequences.  
 SO Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 1.1%; Score 74; DB 1; Length 981;  
 Best Local Similarity 71.0%; Pred. No. 7.1e-07;  
 Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 956 agggatgacattacacactacgaactagcttgcaagatcccaacatgtcttcagag 1015  
 DB 193 AGAGATGCTCTTACGCTTCGCCGAGATTGACAGATCCAGACCAATGCTCCAGAG 252  
 QY 1016 ctggagatccaacctgtgacacctgtgacatgtttcatgtgacatgtaacaatgaaa 1075  
 DB 253 CTGGGATCCAACTCTGTATATCTGTACTGCTTCATGTCACCTGTATACCAAGACAA 312  
 QY 1076 cagtgataaagagtga 1093  
 DB 313 CCGGCTCACCTCGTGTGA 330

## RESULT 13

V06588 ID V06586 standard; cDNA to mRNA; 1106 BP.

AC V06586;  
 DE 03-AUG-1998 (first entry)  
 DE Arabidopsis thaliana SERK LRR homologous EST clone.  
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 KM plant breeding; leucine-rich repeat; ss.  
 OS Arabidopsis thaliana.  
 FH Key Location/Qualifiers  
 FT CDS 142..798  
 FT /tag= a /note= "shows high homology to SERK"  
 FT W09743427-A1.  
 PN 20-NOV-1997.  
 PD 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 DR WPI: 98-086529/08.  
 DR P-PSDB: W47018.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 28; Pages 71-73; 123pp; English.  
 CC The sequence is that of an EST clone showing high homology to  
 CC SERK LRR (leucine-rich repeat) sequences.  
 SO Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;

Query Match 1.1%; Score 74; DB 1; Length 1106;  
 Best Local Similarity 71.0%; Pred. No. 7.4e-07;  
 Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 956 agggatgacattacacactacgaactagcttgcaagatcccaacatgtcttcagag 1015  
 DB 231 AGAGATGCTCTTACGCTTCGCCGAGATTGACAGATCCAGACCAATGCTCCAGAG 290  
 QY 1016 ctggagatccaacctgtgacacctgtgacatgtttcatgtgacatgtaacaatgaaa 1075  
 DB 291 CTGGGATCCAACTCTGTATATCTGTACTGCTTCATGTCACCTGTATACCAAGACAA 350  
 QY 1076 cagtgataaagagtga 1093  
 DB 351 CCGGCTCACCTCGTGTGA 368

## RESULT 14

V06588

ID V06588 standard; cDNA to mRNA; 788 BP.

AC V06588;  
 DE 03-AUG-1998 (first entry)  
 DE Arabidopsis thaliana SERK LRR homologous EST clone.  
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 KM plant breeding; leucine-rich repeat; ss.  
 OS Arabidopsis thaliana.  
 FH Key Location/Qualifiers  
 FT CDS 2..664  
 FT /tag= a /note= "shows high homology to SERK"  
 FT W09743427-A1.  
 PN 20-NOV-1997.  
 PD 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 DR WPI: 98-086529/08.  
 DR P-PSDB: W47020.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 28; Pages 79-80; 123pp; English.  
 CC The sequence is that of an EST clone showing high homology to  
 CC SERK LRR (leucine-rich repeat) sequences.  
 SO Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;

Query Match 1.1%; Score 70.8; DB 1; Length 788;  
 Best Local Similarity 69.6%; Pred. No. 3.2e-06;  
 Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 956 agggatgacattacacactacgaactagcttgcaagatcccaacatgtcttcagag 1015  
 DB 97 AGAGATGCTCTTACGCTTCGCCGAGATTGACAGATCCGAGACCAATGCTCCAGAG 156  
 QY 1016 ctggagatccaacctgtgacacctgtgacatgtttcatgtgacatgtaacaatgaaa 1075  
 DB 157 CTGGGATCCAACTCTGTATATCTGTACTGCTTCATGTCACCTGTATACCAAGACAA 216  
 QY 1076 cagtgataaagagtga 1093  
 DB 217 CCGGCTCACCTCGTGTGA 234

## RESULT 15

V06590 ID V06590 standard; cDNA to mRNA; 1063 BP.

AC V06590;  
 DE 03-AUG-1998 (first entry)  
 DE Arabidopsis thaliana SERK LRR homologous EST clone.  
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 KM plant breeding; leucine-rich repeat; ss.  
 OS Arabidopsis thaliana.  
 FH Key Location/Qualifiers  
 FT CDS 106..762  
 FT /tag= a /note= "shows high homology to SERK"  
 FT W09743427-A1.  
 PN 20-NOV-1997.  
 PD 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 DR WPI: 98-086529/08.  
 DR P-PSDB: W47022.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 28; Pages 86-88; 123pp; English.  
 CC The sequence is that of an EST clone showing high homology to  
 CC SERK LRR (leucine-rich repeat) sequences.  
 SO Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;

Query Match 1.1%; Score 70.8; DB 1; Length 1063;  
 Best Local Similarity 69.6%; Pred. No. 3.5e-06;



Matches	96;	Conservative	0;	Mismatches	42;	Indels	0;	Gaps	0;
QY	956	aggcgatgcatlaacaacttaaggaactagcttgaagaatcccaaatgtcctgcagag	1015						
Db	195	AGGAGATGCTCTTACGCTCTCGCGAGTTTAAACAGATCCGACCATGTTCTCCAGAG	254						
QY	1016	ctgggataccaacccctgtgaaccttgacatggtttcatgtgacatgtaacaatgaana	1075						
Db	255	CTGGGATCCAACTCTGTATCTGTGTAACCTGTGTTCCATGTCACTGTAAACAAAGACA	314						
QY	1076	cagtgtataaagagctga	1093						
Db	315	CCGCGTCACCTCGGTGGA	332						

Search completed: June 23, 2000, 22:48:42  
Job time: 40318 sec



•  
•  
•  
•

• • •

• • • • •

•



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 11:34:14 ; Search time 274.21 Seconds  
(without alignments) 3173.662 Million cell updates/sec

Title: US-09-180-798-1

Perfect score: 6695  
Sequence: 1 tctagatgacgaatcgcgc.....attaatatatttattggtt 6695

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/2/lna/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/2/lna/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/2/lna/5C.COMB.seq:\*  
4: /cgn2\_6/prodata/2/lna/5D.COMB.seq:\*  
5: /cgn2\_6/prodata/2/lna/5E.COMB.seq:\*  
6: /cgn2\_6/prodata/2/lna/5F.COMB.seq:\*  
7: /cgn2\_6/prodata/2/lna/5G.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109.6	1.6	1554	4 US-08-587-680A-24	Sequence 24, Appl
2	61.6	0.9	5733	3 US-08-473-553A-1	Sequence 1, Appl
3	61.2	0.9	2571	1 US-07-717-331F-9	Sequence 9, Appl
4	61.2	0.9	2833	1 US-07-717-331F-1	Sequence 1, Appl
5	61	0.9	2749	1 US-07-717-331F-4	Sequence 4, Appl
6	60.6	0.9	2749	2 US-08-265-628-1	Sequence 14, Appl
7	59.4	0.9	7218	1 US-08-232-463-14	Sequence 13, Appl
8	55.6	0.8	19134	4 US-08-487-826B-13	Sequence 83, Appl
9	54	0.8	4517	6 PCT-US93-06251-83	Sequence 2, Appl
10	53.8	0.8	966	1 US-08-447-185-2	Sequence 7, Appl
11	53.8	0.8	2443	1 US-08-447-185-3	Sequence 3, Appl
12	50.8	0.8	2647	6 PCT-US93-06251-77	Sequence 77, Appl
13	50.8	0.8	2890	1 US-07-928-464-1	Sequence 1, Appl
14	50.8	0.8	2890	6 PCT-US93-07347-1	Sequence 1, Appl
15	50.8	0.8	3033	1 US-08-003-311B-1	Sequence 1, Appl
16	50.8	0.8	3033	1 US-08-261-432-1	Sequence 1, Appl
17	50.4	0.8	1203	2 US-08-602-010A-7	Sequence 7, Appl
18	50.4	0.8	1203	2 US-08-680-726A-7	Sequence 7, Appl
19	50.4	0.8	5495	2 US-08-602-010A-1	Sequence 1, Appl
20	50.4	0.8	5495	2 US-08-602-010A-2	Sequence 2, Appl
21	50.4	0.8	5495	2 US-08-680-726A-1	Sequence 1, Appl
22	50.4	0.8	5495	2 US-08-680-726A-2	Sequence 2, Appl
23	50.4	0.8	10592	2 US-08-680-726A-51	Sequence 51, Appl
24	50.4	0.8	10592	2 US-08-680-726A-52	Sequence 52, Appl
25	47.6	0.7	1804	2 PCT-US93-06251-82	Sequence 82, Appl
26	47.6	0.7	1804	6 PCT-US93-06251-82	Sequence 1, Appl
27	47	0.7	1602	1 US-07-820-011A-1	Sequence 1, Appl

28	47	0.7	1602	6 PCT-US93-00445-1	Sequence 1, Appl
29	46.6	0.7	5852	1 US-07-867-106-2	Sequence 2, Appl
30	46.6	0.7	5852	1 US-07-867-106-2	Sequence 2, Appl
31	46.2	0.7	6768	2 US-08-107-755A-1	Sequence 1, Appl
32	46.2	0.7	8457	1 US-07-991-867B-1	Sequence 1, Appl
33	46.2	0.7	8457	4 US-08-544-332-1	Sequence 1, Appl
34	45.6	0.7	2447	4 US-09-014-969-14	Sequence 14, Appl
35	45.4	0.7	665	4 US-08-883-795A-36	Sequence 36, Appl
36	44.4	0.7	6243	4 US-09-056-075-1	Sequence 1, Appl
37	44	0.7	1574	5 US-09-173-581-12	Sequence 12, Appl
38	44	0.7	3590	1 US-08-587-889-1	Sequence 1, Appl
39	44	0.7	3590	6 PCT-US96-09193-1	Sequence 1, Appl
40	43.8	0.7	1611	1 US-07-820-011A-3	Sequence 3, Appl
41	43.8	0.7	1611	6 PCT-US93-00445-3	Sequence 3, Appl
42	43.4	0.6	3209	1 US-08-105-483-220	Sequence 220, App
43	43.4	0.6	3209	1 US-08-220-151-62	Sequence 62, Appl
44	43.4	0.6	3209	1 US-08-413-118-62	Sequence 62, Appl
45	43.4	0.6	3209	2 US-08-224-391-90	Sequence 90, Appl

#### ALIGNMENTS

RESULT 1  
US-08-587-680A-24  
: Sequence 24, Application US/08587680A  
: Patent No. 5977434  
: GENERAL INFORMATION:  
: APPLICANT: Ronald, Pamela C.  
: APPLICANT: Wang, Guo-Liang  
: APPLICANT: Song, Wen-Yuang  
: APPLICANT: Szabo, Veronique  
: TITLE OF INVENTION: Procedures and Materials for Confering  
: NUMBER OF SEQUENCES: 27  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Townsend and Townsend and Crew LLP  
: STREET: Two Embarcadero Center, Eighth Floor  
: CITY: San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94111-3834  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/587,680A  
: FILING DATE: 17-JAN-1996  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/373,375  
: FILING DATE: 17-JAN-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/475,891  
: FILING DATE: 07-JUN-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 60/004,645  
: FILING DATE: 29-SEP-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/567,375  
: FILING DATE: 04-DEC-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Bastian, Kevin L.  
: REGISTRATION NUMBER: 34,774  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 576-0200  
: TELEFAX: (415) 576-0300  
: INFORMATION FOR SEQ ID NO: 24:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1554 base pairs















```

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-F1s
; US-08-232-463-14

Query Match      0.9%; Score 59.4; DB 1; Length 7218;
Best Local Similarity 1.6%; Pred. No. 0.00023;
Matches 6; Conservative 229; Mismatches 141; Indels 0; Gaps 0;

QY 299 ttctctcgtagcattcttctctctgctatagttacgagtttgcacacatgacg 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1067 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1126

QY 359 ttgaactgtagtctgtctctccatctaagttagtacctgatatagtcctcaact 418
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1127 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1186

QY 419 tcttcgtagccttctcttcgatttcccaagacaagattcttagttaatgatat 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1187 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1246

QY 479 tgctcgggagctgtgctcatttagagatctctactctgtttttaaagagaacaa 538
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1247 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1306

QY 539 acctacttttctctgtctccctttagatcacctgctgagcgcttagactt 598
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1307 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1366

QY 599 taccacctaatactatgtttaccagaacaagctatagtttaccctcccccgcgc 658
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1367 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1426

QY 659 ggaactgagacaaca 674
    : : : : :
Db 1427 yyyyyyygtraccAA 1442

RESULT 8
US-08-487-826B-13/C
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Slim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhuan
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelisen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOC KEY NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; AMTI-SENSE: NO
; US-08-487-826B-13

Query Match      0.8%; Score 55.6; DB 4; Length 19124;
Best Local Similarity 45.9%; Pred. No. 0.0023;
Matches 345; Conservative 0; Mismatches 394; Indels 13; Gaps 4;

QY 1114 ttcttagcaagttacgaataatttactcaattgagcagatgtctcttaataatttctt 1173
    ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 16212 TTTTTCCTAAATAAATAAGACATATATGTATATATATATATATATATATAT 16153

QY 1174 taattcttagctgaagcagacatcatcttaagtatctctactcgaatttaagacataat 1233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16152 ATATGCTTTGTTAAATGATATATATGATATATATGCGTATATATATATATAT 16093

QY 1234 acctttttttaaataatcattagagtggttttccgcagcagcagacatatctttt 1293
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16092 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 16033

QY 1294 ctgtaattcagaacaacctctccgcagataaataataagatttaactcttgac 1353
    ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 16032 TGTGTAATGTGTATGATGATATGAGGTATATATATATATATATATATATATG 15973

QY 1354 taatttttatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1413
    ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 15972 GTATATATATATATATATATATATATATATATATATATATATATATATATAT 15918

QY 1414 tgaagaagatcatctatgaataatttttggcaaccatataattatattcattc 1473
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 15917 TTTCATTAATATTTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15858

QY 1474 tagtataagttaaaaaagctctcatatcaattccaagataaggttca-ttttatagt 1532
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15857 TTTATTTTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15798

QY 1533 ttggagcatcagtttttgaagaatgcaagaatttttgagtttaatttactacca 1592
    || : || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 15797 TTATGTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15741

QY 1593 actttatgcgcagtcagtcgcttgggaatgacacattctcgtgcaacttggcttct 1652
    ||||| : || : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 15740 ATTATTTATTTATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15681

QY 1653 ctggcagcttgaaaaaatttacaactctgtaagacacataacttggaaagcttagtct 1712
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15680 TTTTAAATGTTT---TTTATTTTCTCTTTGTTTATTTTATTTTATTTTATTTTATTT 15625

QY 1713 ttatacagacaatgctttcaatatctgtttaaagtgtaaaagttgaactttagct 1772
    ||||| : | : || : || : || : || : || : || : || : || : || : || : || : ||
Db 15624 TTTATATATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15565

QY 1773 tcagcagtttctcggaataatctatgaagcaacttaaaagctgggcaattttttgta 1832
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15564 AAAATTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15505

QY 1833 ttattcaaatatgttgaattgttactactta 1864

```







```

CORRESPONDENCE ADDRESS:
ADDRESS: Michael L. Goldman
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,185
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/111,078
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman M., Michael L.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1000
TELEFAX: (716)-263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2443 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-447-185-3

Query Match          0.8%; Score 53.8; DB 1; Length 2443;
Best Local Similarity 53.9%; Pred. No. 0.0028;
Matches 159; Conservative 0; Mismatches 127; Indels 9; Gaps 2;

QY 5010 ttctctcgcgaattgcgaagtcgaacgcgaattcttaagtaacca-----tcccttgaa 5063
DB 149 TTCCCTTAGTAGATTGGAGAGACCACTAATTTGATCACAAGTTTAAATTGGAG 208
QY 5064 gagtgagcttgtaagtgatgaaggagcccttgctgtagtgcactctgtgcagttca 5123
DB 209 ATGGGCTTTGGGAGAGGTTTACAGAGGCTGTTGCGTGATGAGCAAGGTGGCCCTGA 268
QY 5124 aaagccttaagaagaacgaacacacagtgcgagctgcagttccaacagaagtgaa 5183
DB 269 AAAGGCTACACCTGAGTCCTCACAAGGTAATGAG---AGTTCGAAACAGAAATGAGA 325
QY 5184 tgattagcagtgctgcatcgaatctctgcgtctacgtgttctgcatacaccta 5243
DB 326 CTCTCTCTTTTGGAGAGATCGCATCTGTTTCATTGATGATGATGATGATGATGATG 385
QY 5244 ccgagcgcctctgtatccatcacatagtggaagtggtgcgtcgtttt 5298
DB 386 ATGAGATATATTCTAATTATAAATACATGAGAAATGGGAACCTCAGAGACATT 440

RESULT 12
PCT-US93-06251-77
Sequence 77, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
```

```

ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-77

Query Match          0.8%; Score 50.8; DB 6; Length 2647;
Best Local Similarity 51.3%; Pred. No. 0.014;
Matches 118; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 5903 aagattatcatcgcgcgaattgaaagctgcaaatatattatggaagaattgagct 5962
DB 1732 AATTATATTCATAGATCTCGCATCGACCAACATTCTAGTGGCAATGACTCATATGC 1791
QY 5963 gctgtagtgatttggttagtgtagctcagtgatgaattacaagataccatgttagcact 6022
DB 1792 AAGATTCTCTACTCGATGCGATGGCCCGATTGATGAAGACAAATGATGATACAGCAAG 1851
QY 6023 gcttaaggagtgatccttgtagcagatgctccgagtagctctgcacgcgaagatcata 6082
DB 1852 GGTCAAGATTCCTCCATCAATGAGACGCGCCCGAGGACGCCCTTACGGGGAGTTACA 1911
QY 6083 gagaagaccatgtcttggtttagtgaataatgctcctagagctaatatc 6132
DB 1912 ATCAAGCTGACGCTGCTTTTGGATCTTACTCAGACAGAGCTGTGTAC 1961

RESULT 13
US-07-928-464-1
Sequence 1, Application US/07928464
Patent No. 5367065
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R.
TITLE OF INVENTION: Constitutive Triple Response Gene and
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz and
ADDRESS: No. 5367065r1s
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
```



```
APPLICATION NUMBER: US/07/928,464
FILING DATE: 19920810
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Suzanne E.
REGISTRATION NUMBER: 32,279
REFERENCE/DOCKET NUMBER: UPN-1086
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2890 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 118..2583
US-07-928-464-1
```

Query Match 0.8%; Score 50.8; DB 1; Length 2890;

Best Local Similarity 52.8%; Pred. No. 0.014; Mismatches 117; Indels 3; Gaps 1;

Matches 134; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

```
QY 5879 aaatgcatgaccattgtatcccaagattatccatcgcatgtaaaagctgcaaatata 5938
DB 2104 AATTATCTTCACAAATGCAATCTCCCAATTGTGCATAGAGATCTAAATCTCCAAACTTA 2163
QY 5939 ttattggaagaagaattgagctgtgtgagtgatttgggttagctagctcattgat 5998
DB 2164 TTGCTTGACAAAATAATACAGTCAAGCTTTGTGATTTGGTCTCCGGGATTGAAGGCC 2223
QY 5999 tacaagatacccatgttaacgactgtgtaaggggtacattggtgacataagctcccgag 6058
DB 2224 AGCACGTTT---CTTCTCTGGAAGTCAGAGCTGGAACCCCGAGTGGATGGCACACAGAA 2280
QY 6059 taccctcgcactggaagatcatcagaagaagaccgatgcttgggttaaggtaatgctc 6118
DB 2281 GTCCTGCGAGATGAGCCGCTAATGAAGAAGTCAGATGTACAGCTTCGGGGTCATCTTG 2340
QY 6119 cttagagctcattac 6132
DB 2341 TGGGAGCTTGCTAC 2354
```

RESULT 14  
PCT-US93-07347-1

```
; Sequence 1, Application PC/TUS9307347
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R.
; TITLE OF INVENTION: Constitutive Triple Response Gene and
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz and
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07347
; FILING DATE: 19930805
; CLASSIFICATION:
```

```
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Suzanne E.
REGISTRATION NUMBER: 32,279
REFERENCE/DOCKET NUMBER: UPN-1086
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2890 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 118..2583
PCT-US93-07347-1
```

Query Match 0.8%; Score 50.8; DB 6; Length 2890;

Best Local Similarity 52.8%; Pred. No. 0.014; Mismatches 117; Indels 3; Gaps 1;

Matches 134; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

```
QY 5879 aaatgcatgaccattgtatcccaagattatccatcgcatgtaaaagctgcaaatata 5938
DB 2104 AATTATCTTCACAAATGCAATCTCCCAATTGTGCATAGAGATCTAAATCTCCAAACTTA 2163
QY 5939 ttattggaagaagaattgagctgtgtgagtgatttgggttagctagctcattgat 5998
DB 2164 TTGCTTGACAAAATAATACAGTCAAGCTTTGTGATTTGGTCTCCGGGATTGAAGGCC 2223
QY 5999 tacaagatacccatgttaacgactgtgtaaggggtacattggtgacataagctcccgag 6058
DB 2224 AGCACGTTT---CTTCTCTGGAAGTCAGAGCTGGAACCCCGAGTGGATGGCACACAGAA 2280
QY 6059 taccctcgcactggaagatcatcagaagaagaccgatgcttgggttaaggtaatgctc 6118
DB 2281 GTCCTGCGAGATGAGCCGCTAATGAAGAAGTCAGATGTACAGCTTCGGGGTCATCTTG 2340
QY 6119 cttagagctcattac 6132
DB 2341 TGGGAGCTTGCTAC 2354
```

RESULT 15  
US-08-003-311B-1

```
; Sequence 1, Application US/08003311B
; Patent No. 544166
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R.
; TITLE OF INVENTION: Constitutive Triple Response Gene
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/003,311B
; FILING DATE: January 12, 1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,464
```



```

; FILING DATE: August 10, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lori Y. Beardsell
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: DPN-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 118..2583
; US-08-003-311B-1

```

```

Query Match      0.8%; Score 50.8; DB 1; Length 3033;
Best Local Similarity 52.8%; Pred. No. 0.015;
Matches 134; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

QY 5879 aaatgcatgacacattgtatcccaagaattatccatcgcatgtaaaagctgcaaatata 5938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2104 AATTATCTTCACATGCGCATCTCCAAATTGTCATAGAGATCTAAATCTCCAAACTTA 2163

QY 5939 ttattggacgaagaatttgaagcgtgttgaagtgatttgggttagctagctcatgcat 5998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2164 TTGGTTGACAAAAAATATACAGTCAAGGTTTGTGATTGTCCTCGCATTTGAAGGCC 2223

QY 5999 tacaagatgacacatttgcagactgctgtaagggttacattggcacatagctccgag 6058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2224 AGCACGTTT---CTTCCCTCGAAGTACGACGCTGGAACTCCGAGTGGATGCGACACAGAA 2280

QY 6059 taacctcgcactggaagatcatcagaagaagacagatgctcttgggttatggataatgctc 6118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2281 GTCCTCGAGATGAGCCGCTAATGAAAGTCAGATGCTGTACAGCTTCGGGCTCATCTTG 2340

QY 6119 ctgaagctcattac 6132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2341 TGGGAGCTTGCTAC 2354

```

Search completed: June 23, 2000, 22:41:28  
 Job time: 40034 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 11:29:27 ; Search time 6198.48 Seconds  
(without alignments)  
4377.901 Million cell updates/sec

Title: US-09-180-798-1  
Perfect score: 6695  
Sequence: 1 tctagatgacgaatcgcg.....aatataattttttatggtt 6695

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202661650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : EST:\*  
1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: em\_est10:\*  
11: em\_est11:\*  
12: em\_est12:\*  
13: em\_est13:\*  
14: em\_est14:\*  
15: em\_est15:\*  
16: em\_est16:\*  
17: em\_est17:\*  
18: em\_est18:\*  
19: em\_est19:\*  
20: gb\_est1:\*  
21: gb\_est2:\*  
22: gb\_est3:\*  
23: gb\_est4:\*  
24: gb\_est5:\*  
25: gb\_est6:\*  
26: gb\_est7:\*  
27: gb\_est8:\*  
28: gb\_est9:\*  
29: gb\_est10:\*  
30: gb\_est11:\*  
31: gb\_est12:\*  
32: gb\_est13:\*  
33: gb\_est14:\*  
34: gb\_est15:\*  
35: gb\_est16:\*  
36: gb\_est17:\*  
37: gb\_est18:\*  
38: gb\_est19:\*  
39: gb\_est20:\*  
40: gb\_est21:\*  
41: gb\_est22:\*  
42: gb\_est23:\*  
43: gb\_est24:\*  
44: gb\_est25:\*

45: gb\_est26:\*  
46: gb\_est27:\*  
47: gb\_est28:\*  
48: gb\_est29:\*  
49: gb\_est30:\*  
50: gb\_est31:\*  
51: gb\_est32:\*  
52: em\_est20:\*  
53: em\_est21:\*  
54: em\_est22:\*  
55: em\_est23:\*  
56: em\_est24:\*  
57: em\_est25:\*  
58: em\_est26:\*  
59: gb\_est33:\*  
60: gb\_est34:\*  
61: gb\_est35:\*  
62: gb\_est36:\*  
63: gb\_est37:\*  
64: gb\_est38:\*  
65: em\_est27:\*  
66: em\_est28:\*  
67: em\_est29:\*  
68: em\_est30:\*  
69: gb\_est39:\*  
70: gb\_est40:\*  
71: gb\_est41:\*  
72: gb\_est42:\*  
73: gb\_est43:\*  
74: gb\_est44:\*  
75: em\_est31:\*  
76: em\_est32:\*  
77: em\_est33:\*  
78: em\_est34:\*  
79: gb\_est45:\*  
80: gb\_est46:\*  
81: gb\_est47:\*  
82: gb\_est48:\*  
83: gb\_est49:\*  
84: gb\_est50:\*  
85: gb\_est51:\*  
86: em\_gss1:\*  
87: em\_gss2:\*  
88: em\_gss3:\*  
89: em\_gss4:\*  
90: gb\_gss5:\*  
91: gb\_gss6:\*  
92: gb\_gss7:\*  
93: gb\_gss8:\*  
94: gb\_gss9:\*  
95: em\_gss5:\*  
96: em\_gss6:\*  
97: em\_gss7:\*  
98: em\_gss8:\*  
99: em\_gss9:\*  
100: em\_gss10:\*  
101: em\_gss11:\*  
102: gb\_gss10:\*  
103: gb\_gss11:\*  
104: em\_gss12:\*  
105: gb\_gss12:\*  
106: gb\_gss13:\*  
107: gb\_gss14:\*  
108: gb\_gss15:\*  
109: gb\_gss16:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query

SUMMARIES



No.	Score	Match	Length	DB	ID	Description
1	211	4.0	532	79	AW233982	AW233982 sf32g05.y
2	245.2	3.7	656	66	AA738547	AA738547 SBRUK5 SO
3	244.8	3.7	1202	84	B09168	B09168 T2EI0-T7 TA
4	207.6	3.1	555	51	AI728030	AI728030 BNLGH1962
5	197	2.9	538	62	AI900110	AI900110 scioIa0.y
6	191.8	2.9	597	51	AI727837	AI727837 BNLGH1920
7	191.8	2.9	703	51	AI729840	AI729840 BNLGH1133
8	184	2.7	471	47	AI486453	AI486453 EST2444774
9	183.8	2.7	380	43	AI165643	AI165643 A087P60u
10	182.6	2.7	475	80	AA329885	AA329885 N201133e
11	178.2	2.7	714	92	AQ0969360	AQ0969360 LERM53TF
12	175.4	2.6	299	29	D25047	D25047 R1CR29756.R
13	173	2.6	464	84	B27171	B27171 T2EI0rF TAM
14	172.2	2.6	991	38	AA738544	AA738544 SBRUK2 SO
15	170.6	2.5	1435	38	AA738545	AA738545 SBRUK3 SO
16	162.2	2.4	555	62	AI896227	AI896227 EST265720
17	160	2.4	383	81	AA4323485	AA4323485 sh67e01.y
18	145.8	2.2	655	46	AA738546	AA738546 SBRUK4 SO
19	143.4	2.1	621	92	AQ0969361	AQ0969361 LERM53TR
20	142	2.1	490	47	AI486325	AI486325 EST244646
21	133	2.0	474	46	AI416897	AI416897 salPeI1.x
22	126.2	1.9	608	79	AW234133	AW234133 sf34f03.y
23	124	1.9	548	51	AI728857	AI728857 BNLGH1118
24	123.2	1.8	647	51	AI730853	AI730853 BNLGH1700
25	120.8	1.8	496	63	AI9966242	AI9966242 701550775
26	119.4	1.8	631	64	AA0667518	AA0667518 660013808
27	118.8	1.8	412	79	AW310334	AW310334 sf34f03.x
28	118.2	1.8	304	35	C22490	C22490 C22490 R1CE
29	117.6	1.8	524	64	AW033961	AW033961 EST277623
30	117.6	1.7	449	35	C22560	C22560 C22560 R1CE
31	112.2	1.7	385	45	AU056335	AU056335 AU056335
32	109.4	1.6	496	64	AW092144	AW092144 EST825240
33	108.4	1.6	572	80	AA2844352	AA2844352 IGL1.275.D
34	106.8	1.6	335	82	CNSO0X10	AA08080670
35	106.6	1.6	428	28	AA08080670	AA08080670 EST012.Su
36	106.6	1.6	438	74	AW201545	AA0801545 sf04d10.y
37	105.6	1.6	438	79	D40033	D40033 R1CST1556A.R
38	105.4	1.6	314	70	AW257240	AW257240 EST105377
39	104.6	1.6	489	80	AAW329840	AAW329840 N201116e
40	104	1.6	525	62	AI897516	AI897516 EST266959
41	102	1.5	459	35	C22359	C22359 C22359 R1CE
42	100.2	1.5	432	63	AI988078	AI988078 sc96e12.y
43	99.6	1.5	592	79	AW234174	AW234174 sf22b02.y
44	98.8	1.5	608	79	CNS010202	AI088559 Arabidops
45	98.6	1.5	287	82	CNS010202	AI088559 Arabidops

## ALIGNMENTS

RESULT	1
LOCUS	AM233982
DEFINITION	AM233982 532 bp mRNA EST 13-DEC-1999
ACCESSION	AF34905.y1
VERSION	GM-cl028
KEYWORDS	5' similar to TR:023921 023921 SOMATIC EMBRYOGENESIS
SOURCE	RECEPTOR-LIKE KINASE. ; mRNA sequence.
ORGANISM	AM233982
	AM233982.1 GI:6566309
	EST.
	soybean.
	Glycine max
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	euphyllophytes; Spermatophytes; Magnoliophyta; eudicotyledons; core
	eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
	Glycine.
REFERENCE	1 (bases 1 to 532)
AUTHORS	Rhoades,R., Keim,P., Vodkin,L., Eppelding,J., Corryell,V., Shamane,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,I., Underwood,K., Stetter,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schwartz,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M.,

**TITLE**  
McGinn, R., Waterston, R. and Wilson, R.  
**JOURNAL**  
Public Soybean EST Project  
**COMMENT**  
Unpublished (1999)  
On Feb 18, 1999 this sequence version replaced gi:4297707.  
Contact: Shoemaker R/Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estevatson.wustl.edu  
This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134 For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or  
info@genomesystems.com web site: www.genomesystems.com  
High quality sequence stop: 403.

**FEATURES**  
source  
1..532  
    location/Qualifiers  
        organism="Glycine max"  
        db\_xref="taxon:3847"  
        clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1713"  
        clone\_lib="Gm-cl028"  
        tissue\_type="roots of 'Supernod' plants"  
        lab\_host="DH10B"  
        note="Vector: pBluescript II XR; site\_1: EcoRI; site\_2:  
XhoI; The mRNA was isolated from roots of Glycine max  
'Supernod' plants generously donated by Dr. Gary Stacey.  
The seedlings were inoculated with Bradyrhizobium  
japonicum, strain USDA110 prior to harvest. Stratigene's  
cDNA synthesis kit (catalog number 200401) was used to  
synthesize the cDNA. First-strand synthesis was performed  
with 5-methyl dCTP, hence the ligated cDNA was  
hemimethylated. A modification of Stratigene's  
first-strand synthesis primer was used. An 'anchor'  
nucleotide (V=A, C, or G) was added to the 3' end of the  
primer [GAGAGAGAGAGAGAGAGAGACTGTCTGACG(T)18V] to anchor  
the primer at the 5' end of the poly(A) tract. After  
second-strand synthesis, the cDNA ends were filled in with  
cloned Pfu DNA polymerase, ligated to EcoRI adapters and  
subsequently phosphorylated. The XhoI site within the  
first-strand synthesis primer was then restricted by  
digestion with XhoI; all XhoI sites in the cDNA would be  
protected by their hemimethylated status. The cDNA  
constructs were size-fractionated with a 500bp cutoff,  
using GIBCOBRL Life Technologies' cDNA Size Fractionation  
column. The column eluent was then ligated into  
Stratigene's pBluescript II XR predigested vector  
(pBluescript II SK(+)) that has been digested with EcoRI  
and XhoI, and phosphorylated by Stratigene). Both the  
white and blue colonies appear to contain recombinant  
plasmids with cDNA inserts, based on size (~25'). This  
library was constructed by Dr. Paul Keim and Dr. Virginia  
Coreyell."

**BASE COUNT**  
145 a 92 c 129 g 165 t 1 others

**ORIGIN**  
Query Match 4.0%; Score 271; DB 79; Length 532;  
Best Local Similarity 79.9%; Pred. No. 1,2e-42;  
Matches 319; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

5800 agaggtcgcgcacatagaacctcccccctgatgtgccacaactcaggagagattgcactgg 5859  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 103 AGAAGCTCAAGATATCCCAACCCCACTTGCTGGCCAGAAAGGAACGATTTGCAATTGGG 162

5860 attcttcaggagccatctcaaatgcatgaccattgcatgcatcccaagattatccatcgca 5919  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 163 ATCTCGAAGGGGCGCTTGCTTATTTCATGATGATCTTGACACCTTAATATTACACCGTGA 222

5920 tgtaaaaagctgcaaatatattattggaagaaatttgaggtctgtgtgattgatttgg 5979  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 223 TGTCAAGACGCAATATATTGTTGAGAGAGATTTGAGCGAGTGTGAGGATTTGG 282







QY	4728	ttctctctctctctctctattatataagaagagacaaatbgtccctcagactatctgt	4787
Db	111	TTAACTATTTTGGTAAATAAATAATTAAGTGGTATATGTCCTCTGAGCCATTTGGG	170
QY	4788	ggggagtagctcgtcgtgctgccttactcttctgtcgcacctgcatagtcattcgtatg	4847
Db	171	GAGAGTGTGCTGTGGTGGCTTACTATTTTGCTGCCCCCTCTCTTACCTTTTGCTTGGT	230
QY	4848	ggcggagagaagaaaccgcgagaaacattctcttgatgctgcaggttagctccgtataatga	4907
Db	231	GGCCTAAAAAAACCTCAAGAAATCTCTTTGATGTTCCCTGGTAGCACATGATCTGC	290
QY	4908	tatcatatgagggcttac-----tgctcgtgactcttgcttcaactgtcatt	4955
Db	291	AATTCACAGCTTTGTTTCACTTTACAGAAATGGATATGATTTATGACTTAAATTTTTCGT	350
QY	4956	agttaacctcagcttgagaggaaccagaaagctgcacctggtcactatgagaagttcttc	5015
Db	351	TCCTGTATGCAGCCGGAAGAGATCCCTGAGGTTCACTTGGGGAGCTTAAAGCGTTCTCTC	410
QY	5016	tgcgagaattgcacaagtcgcagagaaactctttag-----tacacactctggaagagtg	5069
Db	411	TACGGGAATCTTCAAGTANCAACTGATAGCTTCAGCAACAAAAACATTTTGGCCGAGGTG	470
QY	5070	gatttgtaagtgctgataagaaggacgcctctgctgtagtgcctactctgtaagtaaaagc	5129
Db	471	GGTTCGGAAAAAGCTCTCAAAAGGCCGCTCTGTGATGGAACTTTCTCAGTCAAAACGCG	530
QY	5130	ttaaagaagaacagaaacacacagcgtggtgagctgcagttcaaaacagaaatggaatgata	5189
Db	531	TTTAAAGAGAGCCAAACCCACGGTGGGAGCTCCAGTTTATGACAAATGTGAGATGATNA	590
QY	5190	gcatagcctgcatacgaaatctctgcgtcactcgttgcttctgcatag-acaactacagag	5248
Db	591	GCAATGCCGCTTACACAGATATCTCCTCAGCGCTACNTGCGTTCTGTATGAACCTTACGAA	650
QY	5249	cggctctctgtatatcatcatcagtgctaatggaagtgttgctcatgtttaa	5300
Db	651	AGAAATGCTTATGTTATCCTTTACATATGTTAAGGGAATGATGCTTCTGTGTTGA	702

RESULT	4
A1728030	
LOCUS	
DEFINITION	A1728030 555 bp mRNA EST 11-JUN-1999 BMLH19629 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase (Daucus carota), mRNA sequence.
ACCESSION	A1728030
VERSION	A1728030.1 GI:504816
KEYWORDS	EST
SOURCE	upland cotton. Gossypium hirsutum
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE	1 (bases 1 to 555) Blewitt,M., Metz,E.C., Davy,D.F. and Burr,B. ESTs from developing cotton fiber unpublished (1999)
AUTHORS	
TITLE	
JOURNAL	
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3187310.

```

FEATURES
  source      Location/Qualifiers
1. 555       /organism="Gossypium hirsutum"
              /cultivar="Acala Maxxa"

```

```

/ab xref="taxon:3635"
/clone.lib="Six-day Cotton fiber"
/issue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="XLL-Blue"
/notes="Vector: pBluescript II KS+"

```

Query Match	3.18;	Score 207.6;	DB 51;	Length 555;
Best Local Similarity	62.28;	Pred. No. 2.2e-30;		
Matches 432; Conservative	0;	Mismatches 119;	Indels 143;	Gaps 1;

OY	5961	cttctgaagtgatttttgggttagcctgaagctcatatgatattcaaaagatcacccattgtaga	6020
Db	1	CTGTGTTGGTGACTTTGGGTTGGCTTAACCTTATGAGACTATCAAGATACCACTGATACTA	60
OY	6021	ctgtctgaagagggttacacatttggcagacatatagctcccgagtaacctcgcagcttgaaagtcac	6080
Db	61	CTGCTGTATACGTGGGACACAATTTGGACAAATATGCTCTCGTAATCTCTCATACGAAATCTT	120
OY	6081	cagagaagacccgcatgctcttggttatgtggataatgctctcctgagctcattactctggacaga	6140
Db	121	CAGGAAAACCTGATGATGTTTGGGTATGAGTATATCTTTTGGACCTTATTAACGTGGACAG	180
OY	6141	ggagcttttgatctctgctgcgccttcgcagacagatgcatgctatgtctgttttgattggagtat	6200
Db	181	GGGGCTTGTGATCTTGCTGCTGCTGCTTGGCAAAAGATGAGATGATCTTCTGCTGATTG-----	236
OY	6201	gtgtcccgagggttcccttggtttaatttcaacatattagtgcttaactacttgtgtgtg	6260
Db	236	-----	236
OY	6261	gcccttgtttttatttcttcctgcctgctgatttgatttcttactgtatgtatgatatgtacct	6320
Db	236	-----	236
OY	6321	gctttgcaatgctcttttaggttaaaagccctttgaaagagaaaaagtltggagatgctggt	6380
Db	236	-----GGTCAAAAGACTTCTTGAGAGAGAAAAAGCTGGAAATGCTAGT	277
OY	6381	cgatccctgacctgcagaaacaattacatttgaccagaagtlttagcagccttattccaagttagc	6440
Db	278	TGATCTCGATCTCCAAACCAATTAATGTAGTAAACCTGAGGTAAAGAGATTAAATCCAGGTTGC	337
OY	6441	attactctgtaccacagaggttgcgcgaatgtagagcgtcttaagaatgtcacaagtgatgcgaat	6500
Db	338	TCTGCTATGCACACAAAGTTCGCCCAATGAGACCGGCCCAAAAGATGTCAGAAGTGGTTAAAT	397
OY	6501	gcttgaagtgatgtagcctctgcagaaaaagtggagcagtggtgcacaaagtltgaagtcacaa	6560
Db	398	GCTGGAAAGGTGAATGGGTTGGCCGAAAGATGGGATAGTCGCGACGAAAGTGAAGTTCTACG	457
OY	6561	tcaagaacgttgaatattagctccacatcagtaacctctgtaatgtgcttccctagactcgaagaat	6620
Db	458	GCAGAGAGGTGAACCTTCCGCCCTCATCTTATCTCTGATTTGATGTGTGACTCACTACGACAA	517
OY	6621	cttgcatagcttttgaattatctatgltccagaatata	6654
Db	518	TCTGCATGCTGTGATGATTATCCGCTCAGTCAAGGTGA	551
RESULT 5			
LOCUS	AI900110	538 bp	EST
DEFINITION	scol804.y1 Gm-cl012 Glycine max cDNA clone GENOME SYSTEMS CLONE ID		
	Gm-cl012-871.5' similar to TR-023921 023921 SOMATIC EMBRYOGENESIS		
ACCESSION	AI900110		
VERSION	AI900110.1	GI:5606012	
KEYWORDS	EST.		
SOURCE	soybean.		



REFERENCE	TITLE	ORGANISM
1 (Pases 1 to 538)		Glycine max
Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Corvett, V.,		Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Rhanna, A., Bolla, B., Marra, M., Hiller, L., Kneub, T., Martin, J.,		euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core
Beck, C., Wyle, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,		eudicots: Rosidae: eurosids I: Fabales: Fabaceae: Papilionoideae;
Bowers, J., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,		Glycine.
Schurk, R., Ratter, E., Kohn, S., Shin, T., Jackson, I., Cardenas, M.,		
McCann, R., Waterson, R. and Wilson, R.		
Public Soybean EST Project		
Unpublished (1999)		
On Jun 5, 1998 this sequence version replaced gi:3189472.		
Contact: Shoemaker R./Public Soybean EST Project		
Public Soybean EST Project		
Washington University School of Medicine		
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		
Tel: 314 286 1800		
Fax: 314 286 1810		
Email: est@wustl.wustl.edu		
This clone is available through: Genome Systems, Inc. 4633 World		
Parkway Circle St. Louis, Missouri 63134 For further information		
call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314)		
427-3324 or contact: clones@genomesystems.com or		
info@genomesystems.com web site: www.genomesystems.com		
Seq primer: -40RP from gldco		
High quality sequence stop: 411.		
Location/Qualifiers		
1..538		
/organism="Glycine max"		
/db_xref="taxon:3847"		
/clone="GENOME SYSTEMS CLONE ID: Gm-c1012-871"		
/clone_1db="Gm-c1012"		
/tissue_type="Apical shoot tips, 9-10 day old etiolated		
seedlings"		
/lab_host="XL10-Gold"		
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:		
XhoI; This cDNA library was constructed from mRNA isolated		
from the apical shoots of 9 to 10 day old etiolated		
seedlings. The shoot tips including any emerged leaves		
were harvested for mRNA isolation. The cDNA library was		
prepared using the Stratagene pBluescript II XR cDNA		
library construction kit. Complementary DNA was		
synthesized from mRNA using a primer consisting of a poly		
(dT) sequence with a XhoI restriction site. EcoRI adapters		
were ligated to the blunt-ended cDNA fragments followed by		
XhoI digestion. The cDNA fragments were directionally		
cloned into the EcoRI-XhoI restriction site of the		
transformed vector. The ligated cDNA fragments were		
transformed into XL10-Gold host cells. This library was		
constructed by Dr. Randy Shoemaker and Dr. John		
Erpelting."		
BASE COUNT		
153 a 105 c 142 g 138 t		
ORIGIN		
Query Match 2.9% Score 197; DB 62; Length 538;		
Best Local Similarity 73.6%; Pred. No. 2.4e-28;		
Matches 251; Conservative 0; Mismatches 90; Indels 0; Gaps 0;		
6336 tttaggttaaaaccccttttgaagaagaaagtgtgagatgcgtgcagctcgaactgca 6395		
196 ttgggttaaaagacattctttaaagagaaaaaagcttggaaatgttgatcctgatcttaca 255		
6396 gacaattacattgacacagaagttgagcagcttattcaagtagcattacatgtgaccca 6455		
256 AACCAACTATATAGAAACTGAGGTAGAACAGTTAATCCAGGTGGCACTACTCGACACA 315		
6456 ggggttcgccaatggagcgcccttaagatgtccagaagtgatccgaatgcttgaagtgatg 6515		
316 aggtttcccgatgacacgcactaaagatgacgaagtggtgaagaatccttgaagtgatg 375		

QY	6516	cctggagaagaagttggagcagatgcaaaaagtgaagatcatccaacagctgaagt	6575
Db	376	CTTGGCAGAAATAATGGATGATGGCAGAAAGGTGGAAGTCTCCGGCAGAGAGTGAAGCT	435
QY	6576	agctccacatgcgaactctcgaatgatctcgaactgacacagataacttcgaatcttga	6635
Db	436	GGCCCTCATCTCTAATTCATTCGATTTGGATTTGTCGATCAACCGAAATCTGCATCAGTGA	495
QY	6636	attatctgctccaagaataacagcatataaattgataatga	6676
Db	496	GTTATCTGCTCCAGAGTGCCTTGGCAGCAGTATGTAATTTAA	536
RESULT	6		
LOCUS	A1727837		
DEFINITION	A1727837	597 bp mRNA	EST 11-JUN-1999
ACCESSION	U093048	soybean Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase [Daucus carota], mRNA sequence.	
VERSION	A1727837		
KEYWORDS	A1727837.1	GI:5046689	
SOURCE	EST.		
ORGANISM	upland cotton.		
	Gossypium hirsutum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.		
REFERENCE	1 (bases 1 to 597)		
AUTHORS	Blewitt,M., Matz,E.C., Davy,D.F. and Burr,B.		
TITLE	ESTs from developing cotton fiber		
JOURNAL	Unpublished (1999)		
COMMENT	On Jun. 5, 1998 this sequence version replaced gi:3187180.		
	Contact: Ben Burr		
	Biology Department		
	Brookhaven National Laboratory		
	Upton, NY 11973, USA		
	Tel: 516-344-3396		
	Fax: 516-344-3407		
	Email: burrb@nsl.jnl.gov		
	Seq primer: F3 primer		
FEATURES			
Source	Location/Qualifiers		
	1..597		
	/organism="Gossypium hirsutum"		
	/cultivar="Acala Maxxa"		
	/db_xref="taxon:3635"		
	/clone_lib="Six-day Cotton fiber"		
	/tissue_type="immature fiber"		
	/dev_stage="Six days post anthesis"		
	/lab_host="XL1-Blue"		
	/note="Vector: pBluescript II KS+"		
BASE COUNT	165 a 107 c 165 g 148 t	12 others	
ORIGIN			
Query Match	2.9%;	Score 191.8;	DB: 51; Length 597;
Best Local Similarity	69.8%;	Pred. No. 2.5e-27;	
Matches 259;	Conservative 0;	Mismatches 112;	Indels 0; Gaps 0
QY	5828	gattggccaactggagagagatgcacagatcttctagagggcctactcaattgcatt	5887
Db	11	GATTGGCCTCAAGAAAAGAAATACCCCTAGGTGACAGCATGGGGCTGGATACCTTCAT	70
QY	5888	gaccattgtatcccaagattatccatgcgcgcatgtaaaaagctgcacaatatatttgac	5947
Db	71	GAACATTGCACATCCTTAGATCATCTCATCGAGATGGGAAGCGCTGCMAATGTTTATTGAT	130
QY	5948	gaagaatttaagagctgtgttaggtgatttctgggttagctagagctcatgattacaagat	6007
Db	131	GAATATTTTGACAGAGTTGTGGGTGACTTTGGCCTTGGGAAGTTGTGTGATGAGGCGG	190
QY	6008	accacattacagactgctgtaaggggtaccatttggcagacatagctcccgagttactctcg	6067



Db 191 ACAGATGTGACACTCAAGTCTGGGACATGGACACATACACCCGAGTACTTGTCC 250  
 Oy 6068 actggaagatcatcagaagaagacgagatgtcttggtatgagtaatgctccaaagagcc 6127  
 Db 251 ACTGGGAAGTATCAGAAAGAACGAGATGTCTTGATGAGATTATGCTCTAGAGCTT 310  
 Oy 6128 attactgacagagagcctttgatctgcctgccttcgacagatgatgattatgttg 6187  
 Db 311 GTGACAGGTCAACGTGCATGATGATTCTCAGCTTGGAAGATGAGATGTCTGCTT 370  
 Oy 6188 ttgagttgggt 6198  
 Db 371 CTTGACTATGT 381

RESULT 7  
 LOCUS A1729440 703 bp mRNA EST 11-JUN-1999  
 DEFINITION BNGH13383 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase [Daucus carota], mRNA sequence.  
 ACCESSION A1729440  
 VERSION A1729440.1 GI:5048292  
 KEYWORDS EST.  
 SOURCE upland cotton.  
 ORGANISM Gossypium hirsutum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.  
 REFERENCE 1 (bases 1 to 703)  
 AUTHORS Blewitt,M., Matz,E.C., Davy,D.F. and Burr,B.  
 TITLE ESTs from developing cotton fiber  
 JOURNAL Unpublished (1999)  
 COMMENT On Jun 5, 1998 this sequence version replaced gi:3188791.  
 CONTACT: Ben Burr  
 BIOLOGY DEPARTMENT  
 BROOKHAVEN NATIONAL LABORATORY  
 Upton, NY 11973, USA  
 TEL: 516-344-3396  
 FAX: 516-344-3407  
 EMAIL: burrb@nsl.bnl.gov  
 Seq primer: T3 primer

FEATURES  
 source location/Qualifiers  
 1..703  
 /organism="Gossypium hirsutum"  
 /cultivar="Acacia Maxxa"  
 /db\_xref="taxon:3635"  
 /clone\_lib="Six-day Cotton fiber"  
 /issue\_type="Immature fiber"  
 /dev\_stage="six days post anthesis"  
 /lab\_host="XLI-Blue"  
 /note="Vector: pBluescript II KS+"  
 BASE COUNT 188 a 148 c 180 g 186 t 1 others  
 ORIGIN

Query Match 2.9%; Score 191.8; DB 51; Length 703;  
 Best Local Similarity 69.8%; Pred. No. 2.4e-27;  
 Matches 259; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Oy 5828 gatggccaactagagagagatcactagatctctagaggcctatcataatgcat 5887  
 Db 11 GATTGGCCCTCAAGAAAGATAAACCCTAGTGACAGATGTGGCTGAGTACCTTCAT 70  
 Oy 5888 gaccatttgatcccaaatatcatcgcgcatgtaaaagctgaatatatttgagc 5947  
 Db 71 GACATTTGCAATCTTAATCATTCATGAGATGTGAAGCTGCAATGTATATTGGAT 130  
 Oy 5948 gaagatttgagagcgtctgtagatcttggttagctagagctcatgagattacaagat 6007  
 Db 131 GAAGATTGGAAGCAGTGTGCGTACTTTGGCCTTGGAAGTGTGCGATGTGAGCGG 190  
 Oy 6008 acccatgttaagactgtctgtaagggtacacattgggcacataagctccgagtaacctcg 6067

Db 191 ACAGATGTGACACTCAAGTCTGGGACATGGACACATACACCCGAGTACTTGTCC 250  
 Oy 6068 actggaagatcatcagaagaagacgagatgtcttggtatgagtaatgctccaaagagcc 6127  
 Db 251 ACTGGGAAGTATCAGAAAGAACGAGATGTCTTGATGAGATTATGCTCTAGAGCTT 310  
 Oy 6128 attactgacagagagcctttgatctgcctgccttcgacagatgatgattatgttg 6187  
 Db 311 GTGACAGGTCAACGTGCATGATGATTCTCAGCTTGGAAGATGAGATGTCTGCTT 370  
 Oy 6188 ttgagttgggt 6198  
 Db 371 CTTGACTATGT 381

RESULT 8  
 LOCUS A1486453 471 bp mRNA EST 29-JUN-1999  
 DEFINITION EST244774 tomato ovary, TAMU Lycopersicon esculentum cDNA clone CLEB8G2, mRNA sequence.  
 ACCESSION A1486453  
 VERSION A1486453.1 GI:4381824  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 471)  
 AUTHORS Alcalá,J., Vrebalov,J., White,R., Matera,A.L., Viston,T., Holt,I.E., Liang,F., Upton,J., Rongling,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Niernann,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
 TITLE Generation of ESTs from tomato carpel tissue  
 JOURNAL Unpublished (1999)  
 COMMENT On May 7, 1998 this sequence version replaced gi:3121325.  
 CONTACT: David Frisch  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 Jordan Hall, Clemson, SC 29634, USA  
 TEL: 864 656 4366  
 FAX: 864 656 4293  
 EMAIL: dfrisch@CLEMSON.EDU.

FEATURES  
 source location/Qualifiers  
 1..471  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone\_lib="tomato ovary, TAMU"  
 /issue\_type="carpel"  
 /dev\_stage="5 days pre-anthesis to 5 days post-anthesis"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; cDNA - Tomato Carpel EST Library. Oligo(dT)-primed and directionally cloned cDNA in vector Lambda ZAP II with 5' and 3' ends located at the EcoRI and XhoI sites, respectively."  
 BASE COUNT 136 a 88 c 112 g 135 t  
 ORIGIN

Query Match 2.7%; Score 184; DB 47; Length 471;  
 Best Local Similarity 66.8%; Pred. No. 8.1e-26;  
 Matches 262; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Oy 5801 gaggtcagccatcagacactcccccctgattggccaactagagagagattgactaaga 5860  
 Db 54 GAACCTTAACCTGGGAGAGCCTGTTTGAATGTGSCCAACTAGAGACGCTGTGCGGTGGT 113  
 Oy 5861 tctctaggggcctatcataatgcatgacacattgtgacccaagattatcatcgcat 5920



Db 114 ACTGACGCTGACTGATACCTACATACACCTGTAATCCAAAGATTATTCACCGGAT 173  
 QY 5921 gtaaaagctgcaaatatattataggcaagaatttgagctgttgtagtattcgg 5980  
 Db 174 GTTAAAGCAGCATTATGTGTACTAGATGAAGTTTGAAGCTGAGTGGATTTGGC 233  
 QY 5981 ttagctaggtcactgattacaagaataccatgttaccagtctgtgaagggtaccatt 6040  
 Db 234 CTGGCAACCTAGTGTACCTTGAAGAAACCAATGTGACTCAAGTTCGTGTACATG 293  
 QY 6041 gggacacatagctcccgagacactccgcagtgaagtcataagaagaagcagatcttt 6100  
 Db 294 GGCATATAGCTCTCTAATACCTTATCCACTGCAATATCATGAAAAACATGATTTT 353  
 QY 6101 ggtttaggataatgctccttagagctcattactgacagagagctttagatctgtcgc 6160  
 Db 354 GGCTATGGAATCATGCTTTTGGAAATTGTAACCGGCCAACGTGCAATGACTTCTCAGC 413  
 QY 6161 ctggcgaacgatgatgatgttatgttgttga 6192  
 Db 414 CTGAAAGATGAAGATGATGTCTTGTGCTCGA 445

RESULT 9  
 A1165643 380 bp mRNA EST 03-DEC-1998  
 LOCUS A081P60u Hybrid aspen plasmid library Populus tremula x Populus  
 DEFINITION tremuloides cDNA 5', mRNA sequence.  
 ACCESSION A1165643  
 VERSION A1165643 GI:3856928  
 KEYWORDS EST.  
 SOURCE Populus tremula x Populus tremuloides.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 eudicotyledons; Magnoliophyta; eudicotyledons; core  
 eudicotyledons; Rosidae; eurosids I; Malpighiales; Salignaceae; Populus.  
 1 (bases 1 to 380)  
 Sterky,F., Regan,S., Karlsson,J., Hertzberg,M., Rohde,A.,  
 Holmberg,A., Amlin,B., Bhalerao,R., Larsson,M., Villarroel,R., Van  
 Montagu,M., Sandberg,G., Olsson,O., Teerl,T.T., Boerjan,W.,  
 Gustafsson,P., Uhlen,M., Sundberg,B. and Lundberg,U.  
 Gene discovery in the wood-forming tissues of Poplar: Analysis of  
 5,692 expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)  
 99007314  
 On Jan 17, 1998 this sequence version replaced gi:2044473.  
 Contact: Sterky F  
 Department of Biotechnology  
 Royal Institute of Technology (KTH)  
 Teknikringen 34, S-100 44 STOCKHOLM, Sweden  
 Tel: +46 8 790 8287  
 Fax: +46 8 24 54 52  
 Email: frederik@biotech.kth.se  
 PCR Primers  
 FORWARD: AAGGGGATGTCGTGCAAGCG  
 BACKWARD: GCTCCGCTGTGATGTGTG  
 Seq primer: CCGTGAACGACGCGCAG  
 High quality sequence stop: 380.  
 Location/Qualifiers  
 1..380  
 /organism="Populus tremula x Populus tremuloides"  
 /db\_xref="taxon:47664"  
 /clone\_lib="Hybrid aspen plasmid library"  
 /tissue\_type="Cambial region"  
 /dev\_stage="1.5 m actively growing tree"  
 /lab\_host="E.coli"  
 /note="Vector: pBluescript SK; Site.1: SalI; Site.2: NotI;  
 Cambial region tissues, including developing xylem, the  
 meristematic cambial zone and the developing and mature  
 phloem, was harvested from 1.5 m actively growing trees.  
 cDNA was prepared and cloned into lambda gt22a. DNA was  
 isolated and subcloned into pBluescript SK using SalI and

BASE COUNT 109 a 67 c 99 g 102 t 3 others  
 ORIGIN  
 Query Match 2.7%; Score 183.8; DB 43; Length 380;  
 Best Local Similarity 80.1%; Pred. No. 8.9e-26;  
 Matches 226; Conservative 0; Mismatches 55; Indels 1; Gaps 1;  
 QY 5918 gatgaaagctgcaaatatattataggcaagaatttgagctgttgtagtattc 5977  
 Db 1 GATGTTAAAGCTGCAATATTTGTGTGACAGAGAAATTCAGGWTGATGCGGGATTTT 60  
 QY 5978 gggtagctaggtcactgattacaagaataccatgttaccagtctgtgaagggtacc 6037  
 Db 61 GGGCGNCAATATATNACTACAGAGATNCGATGTCCACTGCTGT-CGTGGTACA 119  
 QY 6038 attggacacatagctcccgagacactccgcagtgaagtcatacagaagaagcagatgc 6097  
 Db 120 ATAGGACACATGCGTCCAGAGTACTTACTGTAATCATCTGAAATCTGAAACATGATGT 179  
 QY 6098 ttggtttaggataatgctccttagagctcattactgacagagagctttagatctgtc 6157  
 Db 180 TTTGGTATGGAATCATGCTTTTGGAAATTGTAACCGGCCAACGTGCAATGACTTCTCAGC 239  
 QY 6158 cgccttgcaacgatgatgatgttatgttgttga 6199  
 Db 240 CGGCTGGCAATGATGACGATGATGATTTACTTGAATGGGA 281

RESULT 10  
 A329855 475 bp mRNA EST 28-JAN-2000  
 LOCUS A329855  
 DEFINITION N20113ze rootphos(-) Medicago truncatula cDNA clone MHRP-24H06,  
 mRNA sequence.  
 ACCESSION A329855  
 VERSION A329855.1 GI:6800450  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 eudicotyledons; Magnoliophyta; eudicotyledons; core  
 eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 Medicago.  
 1 (bases 1 to 475)  
 Harrison,M.J., Liu,J., Gonzales,M.B. and Ellis,L.  
 ESTs from phosphate starved roots  
 Unpublished (1999)  
 On May 1, 1997 this sequence version replaced gi:2059389.  
 Contact: Maria J. Harrison  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA  
 Tel: 580-223-5810  
 Fax: 580-221-7380  
 Email: mjharrison@noble.org  
 Developmental stage: phosphate starved; Date: 01/23/00; Submitted  
 to the Database of Expressed Sequence Tags (dbEST) on 01/26/00;  
 More information is available at  
 'http://chrysis.tamu.edu/medicago'.  
 Seq primer: 73.  
 Location/Qualifiers  
 1..475  
 /organism="Medicago truncatula"  
 /cultivar="genotype A17"  
 /db\_xref="taxon:3880"  
 /clone="MHRP-24H06"  
 /clone\_lib="rootphos(-)"  
 /tissue\_type="roots"  
 /note="Vector: pBluescript SK -; Site.1: EcoRI; Site.2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA. The  
 cDNA was directionally ligated into the Unizap XR vector  
 from Stratagene and packaged using GigaPack III Gold

#### FEATURES

source  
 /organism="Populus tremula x Populus tremuloides"  
 /db\_xref="taxon:47664"  
 /clone\_lib="Hybrid aspen plasmid library"  
 /tissue\_type="Cambial region"  
 /dev\_stage="1.5 m actively growing tree"  
 /lab\_host="E.coli"  
 /note="Vector: pBluescript SK; Site.1: SalI; Site.2: NotI;  
 Cambial region tissues, including developing xylem, the  
 meristematic cambial zone and the developing and mature  
 phloem, was harvested from 1.5 m actively growing trees.  
 cDNA was prepared and cloned into lambda gt22a. DNA was  
 isolated and subcloned into pBluescript SK using SalI and







FEATURES	High quality sequence stop: 464.
source	Location/Qualifiers
	1. 464
	/organism="Arabidopsis thaliana"
	/strain="Columbia"
	/db_xref="taxon:3702"
	/clone="T2510"
	/clone_1ib="TAMU"
	/sex="hermaphrodite"
	/note="Vector: BeloBACII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"
BASE COUNT	119 a 87 c 115 g 143 t
ORIGIN	

[illegible]



Contact: Stockhaus J  
Institut fuer Entwicklungsbiologie und Molekularbiologie der  
Pflanzen  
Reinrich-Heine-Universitaet  
Universitaetsstrasse 1, 40225 Duesseldorf, Germany  
Tel: (49) 211-81-14953  
Fax: (49) 211-81-14871  
Email: stockh@uni-duesseldorf.de

Insert Length: 1500    Std Error: 0.00

**FEATURES**  
**source**

1.991

```
/organism="Sorghum bicolor"  
/strain="cv. TX430"
```

BASE COUNT	253 a	194 c	293 g	25
ORIGIN				

Query Match	2.68;	Score 172.2;	DB 38;	Length 991
Best Local Similarity	68.08;	Pred. No. 1.5e-23;		

Oy 5828 gaagggcaactggagagagatgcacagatctcttlaagggcctatcctaattgcac 5887  
 Db 358 gatggctcttcaagaaagaaagcagtgccgtatnagccacagctcctggactggaattttgcac 417  
 Oy 5888 gaccattgcatcccaagatlabcatcgcgcgtatglaaaagcttgaanaabataatttgcac 5947  
 Db 418 gagcactgcacatccttagattatctcgtacgcgtcaagagctgcacaaatgcttcttgat 477  
 Oy 5948 gaagaaattgaagcctttgtaagtgaattttggttagtcgaagcctagatataaagat 6007  
 Db 478 gaagcttttaaccggcttctgttgatcttccgcttggcccaacgtcgtagatgtaacagaag 537  
 Oy 6008 acccaatctacgactgcgtglaaagggctaccatttggcacatagctcccgatcctctcg 6067  
 Db 538 acatctgtgactactcaggtccctggaaamaagggctgcacatgtgccctcgaaagattttgcc 597  
 Oy 6068 actggaagcatcacaagaagaacgcgtgctcttggttttagataatgctctcctagagctc 6127  
 Db 598 actgggaagtcataccagaagaacacagatgtttttggttgcagatataatgctcttcagacta 657  
 Oy 6128 ataccctggaagagagccttttgaattctctgcgccttcgaaagatgatgatgt 6180  
 Db 658 gtccactgtatagcgaagccatagactttttacgtcttggaggaagaagatgagatgt 710

RESULT 15  
AA738545

LOCUS	1435 bp	EST	17-AUG-1998
DEFINITION	SbBRAK3 Sorghum bicolor cv. TX430	leaf	Sorghum bicolor cDNA clone

ACCESSION	AA738545
VERSION	AA738545.1
KEYWORDS	GI:3421470 EST.

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (pages 1 to 1435)	Annen, F., Chang, J.-I., Paterson, A. H. and Stockhaus, J.	Characterization of 14 different putative protein kinase cDNAs of the C4 plant <i>Sorghum bicolor</i>	Mol. Gen. Genet. 259 (1), 115-122 (1998)

MEDLINE 98409267  
COMMENT On Nov 29, 1993 this sequence version replaced g1:636085

Institut fuer Entwicklungsbiologie und Molekularbiologie der  
Pflanzen  
Heinrich-Heine-Universitaet  
Universitaetsstrasse 1, 40225 Duesseldorf, Germany  
Tel: (49)-211-81-14953  
Fax: (49)-211-81-14871  
Email: stockhausen@duesseldorf.de

Insert Length: 1900    Std Error: 0.00

**FEATURES**  
**SOURCE**

1. .1435

```
/organism="Sorghum bicolor"  
/strain="cv. TX430"
```

BASE COUNT	365 a	300 c	389 g	381 t
ORIGIN				

```
Query Match      2.5%; Score 170.6; DB 38, Length 1435;
Best Local Similarity 67.7%; Pred. No. 2.9e-23;
Matches 239; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
```

OY	5828	gattggcccaactcagggaaagagattgttcacacagatctcttaagggccctcaataatgat	5887
Db	802	gATTGGTTCtGCAAGAAAGCAGAGGGCGATTAGGACAGAGCTCGTGGACtGAGATTTTGCAc	861
OY	5888	gaccatttgcatacccaagatatccatcgcgatgtaaagctgcacaatatattatgtgac	5947
Db	862	GACCACTGGCAACCTTAAATATTACATCGTCGACGTCGAAGGGCTGGCAATTGGTCTGTGAT	921
OY	5948	gaagaatttcgaagctggtttgtatgtagtatttggtatagctacatgcgatatacaagat	6007
Db	922	GAAgGTTTtTGAAACCGGTGTtTGAGATTTCGGCTTGCGCAAGCTGtGGATtTGACGAAG	981
OY	6008	accacatgttaagcactgcctctgtataaggggtacacattggtgcacatagtctccgaagtacctctg	6067
Db	982	ACATCTGTGGCTGCTCAAGGTCTCGGAACATAGGGGTCAATTGGCCCTGGAAATTTTGTCC	1041
OY	6068	actggaagaatctacagaaagacccagatctctttgtttatgtgataatgcctcctagaagtc	6127
Db	1042	ACtGGGAAGtCATCTCGAAGAAAGACGAAGTtTTTGGTtACGGCATTAAAGCTtTCGAGCTA	1101
OY	6128	attactcgcagacagggctcttttgcatactgcctctgcctctgtgcgaacgatgatgtatgt	6180
Db	1102	GTCACtGGtGCACGtGCCATtGACtTTTtTACGtCTGGAGAAAGtGATtGTG	1154

Search completed: June 23, 2000, 19:05:52  
Job time: 27385 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 23:30:33 ; Search time 12463.9 Seconds

(without alignments)  
-141.659 Million cell updates/sec

Title: US-09-180-798-2

Sequence: 1815  
1 gacaaataccatggaatat.....ttaaaaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY\_NMC

Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

Geneml:\*  
1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_om:\*  
4: gb\_ov:\*  
5: gb\_pat:\*  
6: gb\_ph:\*  
7: gb\_p11:\*  
8: gb\_p12:\*  
9: gb\_p12:\*  
10: gb\_p12:\*  
11: gb\_p12:\*  
12: gb\_p12:\*  
13: gb\_p12:\*  
14: gb\_p12:\*  
15: gb\_p12:\*  
16: gb\_p12:\*  
17: gb\_p12:\*  
18: gb\_p12:\*  
19: gb\_p12:\*  
20: gb\_p12:\*  
21: gb\_p12:\*  
22: gb\_p12:\*  
23: gb\_p12:\*  
24: gb\_p12:\*  
25: gb\_p12:\*  
26: gb\_p12:\*  
27: gb\_p12:\*  
28: gb\_p12:\*  
29: gb\_p12:\*  
30: gb\_p12:\*  
31: gb\_p12:\*  
32: gb\_p12:\*  
33: gb\_p12:\*  
34: gb\_p12:\*  
35: gb\_p12:\*  
36: gb\_p12:\*  
37: gb\_p12:\*  
38: gb\_p12:\*  
39: gb\_p12:\*  
40: gb\_p12:\*  
41: gb\_p12:\*  
42: gb\_p12:\*  
43: gb\_p12:\*  
44: gb\_p12:\*

45: gb\_htg7:\*  
46: em\_htg1:\*  
47: em\_htg2:\*  
48: em\_htg3:\*  
49: em\_htg4:\*  
50: em\_htg5:\*  
51: gb\_p13:\*  
52: gb\_p13:\*  
53: gb\_p13:\*  
54: gb\_htg8:\*  
55: gb\_htg9:\*  
56: gb\_htg10:\*  
57: gb\_htg11:\*  
58: gb\_htg12:\*  
59: gb\_htg13:\*  
60: gb\_htg14:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1815	100.0	5 A67797	A67797 Sequence 2
2	1755	96.7	8 DCU93048	U93048 Daucus caro
3	934.8	51.5	5 A67827	A67827 Sequence 32
4	560.2	30.5	5 A67796	A67796 Sequence 1
5	434	23.9	8 F23M19	AC007454 Arabidops
6	398	21.9	7 ATEF17M5	AL035678 Arabidops
7	395.6	21.8	8 AC012654	AC012654 Arabidops
8	372.4	20.5	50 AC006436	AC006436 Arabidops
9	317.6	17.5	8 SBRXK1	Y14600 Sorghum bic
10	270.8	14.9	5 A67815	A67815 Sequence 20
11	252.6	13.9	8 AC013395	AB013395 Arabidops
12	246.8	13.6	7 AB025639	AB025639 Arabidops
13	216.2	11.9	8 F17123	AF160182 Arabidops
14	215.4	11.9	8 AC002292	AC002292 Genomic s
15	209.4	11.5	50 AC005170	AC005170 Arabidops
16	203	11.2	7 AB005234	AB005234 Arabidops
17	200.8	11.1	7 AB012245	AB012245 Arabidops
18	178.4	9.8	7 AB007644	AB007644 Arabidops
19	164.6	9.1	8 AC002330	AC002330 Arabidops
20	163.8	9.0	50 AC005312	AC005312 Arabidops
21	161.4	8.9	7 RICEPK1	L27821 Oryza sativ
22	160.8	8.9	8 ATEF884	AL034567 Arabidops
23	160.8	8.9	7 ATEF10M6	AF024648 Arabidops
24	158.4	8.7	8 ZMA01016	AF024648 Arabidops
25	157.8	8.7	8 ZMA01016	AF024648 Arabidops
26	153.8	8.5	8 ZMA01016	AF024648 Arabidops
27	153.6	8.5	8 ZMA01016	AF024648 Arabidops
28	153.2	8.4	8 ATEF884	AL034567 Arabidops
29	153.2	8.4	8 ATEF884	AL034567 Arabidops
30	153.2	8.4	8 ATEF884	AL034567 Arabidops
31	153.2	8.4	8 ATEF884	AL034567 Arabidops
32	149	8.2	8 ATEF884	AL034567 Arabidops
33	145.6	8.0	8 ATEF884	AL034567 Arabidops
34	145.6	8.0	8 ATEF884	AL034567 Arabidops
35	138.4	7.6	5 ATEF884	AL034567 Arabidops
36	138.4	7.6	5 ATEF884	AL034567 Arabidops
37	138.4	7.6	5 ATEF884	AL034567 Arabidops
38	133.8	7.4	8 ATEF884	AL034567 Arabidops
39	133	7.3	7 ATEF884	AL034567 Arabidops
40	131.4	7.2	50 AC004684	AB016890 Arabidops
41	130.6	7.2	8 AC002392	AC002392 Arabidops
42	129.8	7.2	7 ATEF884	AL034567 Arabidops
43	128.2	7.1	8 ATEF884	AL034567 Arabidops
44	128.2	7.1	8 ATEF884	AL034567 Arabidops
45	128.2	7.1	8 ATEF884	AL034567 Arabidops

#### ALIGNMENTS



RESULT 1  
LOCUS A67797 1815 bp DNA PAT 05-MAY-1999  
DEFINITION Sequence 2 from Patent WO9743427.  
ACCESSION A67797  
VERSION A67797.1 GI:4756623  
KEYWORDS  
SOURCE  
ORGANISM  
Daucus carota  
Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta:  
euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core  
eudicots: Asterales; easterids II; Apiales; Apiales; Daucus.  
REFERENCE  
1 (bases 1 to 1815)  
De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE  
PRODUCTION OF APOMICRITIC SEED  
JOURNAL  
Patent: WO 9743427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
Source location/Qualifiers  
1..1815  
/organism="Daucus carota"  
/db\_xref="taxon:4039"  
94..1755  
/note="unamed protein product"  
/codon\_start=1  
/protein\_id="CABA2241.1"  
/db\_xref="GI:4756624"  
/translation="MNRNSINILNTYQFTDAYLDKYGVLTLELYSNISGPIPSDLG  
NLTNLSLDLYMNSFSGIPDITLKLRLFLNLNNSLSGPIPSLNTITLQVLDL  
SNRSLSPVDPNCSFSLFTPISEFANLNLCGSPVRCPCSPSPSPSPSPSPSP  
PONGPTGALAGVAAAGALLFAPAPAFAMMRBRKREHFVPAEEDVEVLGOLK  
RFSRLFOVATDFSTILRGSGEKVYKGLADGSLVAVYRLKEBRPGGELOFOREV  
EMISAVHNRILRLRRCMTPERLIYVPMANGSVASCIREOSPSPSPSPSPSPSP  
ALGSRKSLSTIHDCKPIIHRDVKANIIIDEFPAVYVDEGLARLMDKDHHTTA  
VRGTIGTAPETLSTGKSSSEKTDVFGIMLELTGQRAFDLARLNDVDVMLDLY  
KSLRKERLEMLVDPDENNYIDTEVBOILQVALLCTGSGSPMRPRKSEVYVRLLEDG  
LAEKDEOKVEVTHQVDELAPHRTESEMILSDNHLAEFLSGPR"  
BASE COUNT 531 a 354 c 415 g 515 t  
ORIGIN

Query Match 100.0%; Score 1815; DB 5; Length 1815;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1815; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gcaaatccattgaataattgaacctaataattagtagtgcaggttaattcaaa 60  
DB 1 GCAAAATACCATTTGAATTTGAACCTAATTAATTAGTAGTCAAGGTTAAATCANA 60  
QY 61 ccatttaatttactttaaaaaaattctatacgaaatcgtaacagataataatla 120  
DB 61 CTCATTATTTTACTTTAAAAAATAATTCTATATGAATCGAAGATATAATATA 120  
QY 121 aattacatgcagttcaatgcagttcaatgcagaaatgcaggtctctatgcattgag 180  
DB 121 AATTACATGCAATTCACGATGCTTACCTTGCAAAATAGGGGCTTATATACATGGAG 180  
QY 181 cttaacaagaataaataagtgacaaattcctagtagtctctggaaatcagaaattg 240  
DB 181 CTTTACAGCAATTAAGTAGTGACCAATTCCTATGATCTTGGGAATCTGCAAAATTG 240  
QY 241 gtgagcttgacatatacagaaatgcagaaattcctagtagtctctggaaatcagaaattg 300  
DB 241 GTGAGCTTGACCTATACATGATAGCTCTCTGAGACCTACGACACATTAGGAAG 300  
QY 301 cttaacaagaataagttctgtgctctcaacaacaagaagctctctgtgcgaattccaatg 360  
DB 301 CTTACAGAGCTAAGATTCTGCTCTCAACAACAAGCCTCTGCTGTCCAATTCCAATG 360  
QY 361 tcaactacataattacaactctcaagtcctggatttacaacaacatcggtatcaga 420  
DB 361 TCACCTACATATATTCACACTCTTCAAGTCTGATTTATCAAAACAATCGGCTATACAGA 420

QY 421 ccagtaaccggaataatggtcattcttctgtttacacctatcagatttgcacaaatttg 480  
DB 421 CCAGTACCGGAATAAGGCTCATTTCTTTGTTACACCTATCACTTTGGCAATAATTG 480  
QY 481 aattatctgagcccgtaacatcgagagcccgccctgatactcccaatttgcacaa 540  
DB 481 AATTATATGAGACCCGTAACCTGGAGGCGCTGCCCTGATGTCCCATTTTGGCACCA 540  
QY 541 cctccgttccatccacatcaacagtaacagcccccagaaatgltcccaatgagct 600  
DB 541 CCTCCGTTTCATCCACCATCAACAGTACAGCTCCAGACAAAATGGTCCACAGGAGCT 600  
QY 601 attgcttgagagatgagtcgctgctgctgcttactggttctgtaacatgagcaatt 660  
DB 601 ATTGCTGGGGAGTAGCTGCTGCTGCTGCTTACTGTTGCTGCACTCGCAATGGCAATT 660  
QY 661 gtagtctgagagagaaacccgagaaacattcttgaatgctgagcagctgagagagac 720  
DB 661 GCATGCTGCGGAGAAAGAAACCGGAGACATTTCTTGTATGTGCCAGCTGAAGAGGAC 720  
QY 721 ccagaatgacatctggtcgaactgaagaagttctctctgagagaattgcgaatgcaag 780  
DB 721 CCAGAAGTGCACCTTGCTGCAACTGAAAGGCTTCTCTGCGAGAAATGCAAGTGGCAAG 780  
QY 781 gatacttttagtaccatacttggagagagtgattgtgtaagtgatgaagagcgact 840  
DB 781 GATACCTTTTAGTACATACCTTGGAGAGGTGATTTGGTAAAGGTGATTAAGGAGCGCTT 840  
QY 841 gctgatgctcaactgttagcagttaaagagcttaagaagaacgaacacaggtgtgtgag 900  
DB 841 GCTGATGCGCTCCTGCTGAGCAGTAAAGGCTTAAAGGCTTAAAGAACACACACCGAGGTGAG 900  
QY 901 ctgagatctcaacagagagtggaatgtagtgcgtgctgctgagaaatctctcgt 960  
DB 901 CTGCAATTTCAACAGAGGTGGAATGATTCAGAGGCTGCTGCAATCTTCTGCGT 960  
QY 961 ctacgtgttctctgatacacaacaaacagagcgctctctgtatataccaatagtgctaat 1020  
DB 961 CTACGTGTTTGTGATGATGACCAACAGAGGCGCTTCTTATATTCATACATGAGCTTAT 1020  
QY 1021 ggaagtgttgctgctgtttaaagagacgtaagccatagaactccctgtatggca 1080  
DB 1021 GGAAGTGTGCTGCTGCTGTTTAAAGAGCGTCAAGCATAGAACCTCCCTGATTGGCA 1080  
QY 1081 actaagaagaagattgcacaaatgcagaaatgcagaaatgcagaaatgcagaaatgc 1140  
DB 1081 ACTAGGAAGAGATTGCACTAGGATCTGTAAGGCGCTTCTTATTTGCAAGACATGCT 1140  
QY 1141 gatcccaagattatccatcgtatgtaaaagctgcaaatatattatggacgaagaatt 1200  
DB 1141 GATCCCAAGATTATCCATGCTGATTAAGCTGCAAAATATATTATGAGAGGAATTT 1200  
QY 1201 gaggctgttgatggtgatttggttagtagagcctgaggtttcaagaagatacgaatgt 1260  
DB 1201 GAGGCTGTGTAGGTGATTTGGGTTAGCTAGGCTCAATGATTAAGGATTAACCAATGT 1260  
QY 1261 acaactgctgaagaggtacatttggtcacaatagctccagagttacatctgcagtgaaag 1320  
DB 1261 ACAACTGCTGAAGGGGTACCTTTGGGCTACATAGTCCCGAATACCTCTCGACTGGAAAG 1320  
QY 1321 tcaatgaagaagacgaatgcttctgttgaatgagtaatgcttgaagctcaatcagta 1380  
DB 1321 TCATCAGAGAGACCCGATGCTTGTGTTATGGGATTAAGCTCTTGAACCTATTAAGTGA 1380  
QY 1381 cagaagagctttgatactgctgagcttgagagagatgagatgattatgttttgattgg 1440  
DB 1381 CAGAGAGCTTTTATGATTTGCTGCGCTTGGCAAGATGATGATGATGATGATGATGATG 1440  
QY 1441 gttaaaagccttttgaagaagaagaagttgagatgctgctgacatcagctgagagagac 1500  
DB 1441 GTTAAAAGCCTTTTGAAGAAGAAAAGTTGGAGATGCTGATGATGATGATGATGATGATG 1500  
QY 1501 aattacattgacacagaagttgagcagcttattcaagtagcattctctgtacccaaggt 1560



|||||  
Db 1501 AATTACATTGACAGAAAGTTGAGCAGCTTATTCAGTAGCATCTGTGATCCAGGGT 1560  
QY 1501 tcggcaatgagagcggttaagatgctcagagtgatccgaatgcttaagtgatgctt 1620  
Db 1501 TCGCCAATGAGCGCGCTTAAGATGTCAGAGTAGTCCGAATGCTTAAGATATGACCTT 1620  
QY 1621 gcagaaagtgagagcgagtgcaaaaagtagaagtcacatccatcaagcagaatagct 1680  
Db 1621 GCAGAAAAGTGGAGCGAGTGGCAAAAAGTAGAATCATCATCAAGCAGTAGATAGCT 1680  
QY 1681 ccaatcgaatcttgatgagatccctagactcgacagataactgcatgcttttaata 1740  
Db 1681 CCACATCGAATCTTGATGAGATCCTAGACTCGACAGATAACTGTCATGCTTTGAAATTA 1740  
QY 1741 tctgtccaaagataacagcatataaagtgaatgaatatttttttttttttttttttt 1800  
Db 1741 TCTGTCCAAAGATTAACAGCATATAAATGTAATGTAATTTTATTTTATGCTTAA 1800  
QY 1801 aaaaaaaaaaaaaa 1815  
Db 1801 AAAAAAAAAAAAAA 1815  
  
RESULT 2  
LOCUS DCU93048 1755 bp mRNA PLN 28-JUN-1997  
DEFINITION Daucus carota somatic embryogenesis receptor-like kinase mRNA,  
complete cds.  
ACCESSION U93048  
VERSION U93048.1 GI:2224910  
KEYWORDS  
SOURCE  
ORGANISM  
Daucus carota  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Asteridae; euasterids II; Apiales; Apiales; Daucus.  
REFERENCE  
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.  
TITLE A leucine-rich repeat containing receptor-like kinase marks somatic  
plant cells competent to form embryos  
JOURNAL Development 124 (10), 2049-2062 (1997)  
MEDLINE 97313247  
REFERENCE  
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.  
TITLE Direct Submission  
JOURNAL Submitted (12-MAR-1997) Molecular Biology, Agricultural University  
of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands  
FEATURES  
SOURCE  
Location/Qualifiers  
1..1755  
/organism="Daucus carota"  
/db\_xref="taxon:4039"  
34..1695  
/note="SERK"  
/codon\_start=-1  
/product="somatic embryogenesis receptor-like kinase"  
/protein\_id="AA061708.1"  
/db\_xref="GI:2224911"  
/translation="NNRNSINILNQFIDAVLDKRYVLMLELXSNNSISPIPSDLG  
NLTVLSIDLYANSFSGIPDITGLTRLRFLNNNSLSGIPMLSNITTLAYLD  
SNRRLSGPVNDGSLFTPISEFANLNLNCGPVGPCGSPFSPPEPIPSYVOP  
PGONGPTGAIAGVAGAAALFEAPAAAFAMRRRRRPRHFDVAEDPEVHLGOLK  
FSLRLQVATDPTFSTILRGGGKRYKRIADGSLVANKRKERTGGELQRTQY  
EMTSMAYHRLILRQPCMPTRRLVRYRYANGSASLRLRQSEPLDPFRKRI  
ALGSAKSLSLHDHCPKILIRHYKANLILDEFAVAVDGLRLNDYKDTHTTA  
VRGLGTIAEPISTKSSKEDIVFGLMELDELTGGAFAFLARLADDVMLDWY  
KSLKERKLEMLVDPDENNYIDTVEEOLIOVALICTGSPMERKMSBYVMLEGG  
LAERKMDQMEVFIHQDVELAPRHSWILSDTDLHAFLSGPR"  
BASE COUNT 506 a 347 c 407 g 495 t  
ORIGIN  
Query Match 96.7%; Score 1755; DB 8; Length 1755;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1755; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 cttatttaatttactcttaaaaaataattctatatgaaatcgttaacagatataatttta 120  
Db 1 CTTATTAAATTTTACTTTAAAAAATTAATTCATATGAATCTTAACAGATTAATTAATTA 60  
QY 121 aataacatgagctacatgctatgcttaacctgaacaataatggggtctctatgactgag 180  
Db 61 AATTACATGAGTTCCTATGCTATGCTTACCTTGACAAATATGGGGCTTCTTATACATGGAG 120  
QY 181 cttaacagcaataacaaagtgaccgaatccctagtgactcttggaactcgtacaatttg 240  
Db 121 CTTTACAGCATTAATTAAGTGGACCAATTCCTAGTGAATCTGGGAATCTTACAAATTTG 180  
QY 241 gtagactgagctatatacagataagctctctctgtgacctataccggacaacttgagaag 300  
Db 181 GTAGCTTGACCTATACATGATGATAGCTTCTGTGACCTATACCGACACTTATGGAAG 240  
QY 301 cttaacagctaaagattcttgctctcaacaacaacagccctctgtgccaaatcccaatg 360  
Db 241 CTTACAAAGCTAAGATTTCTGCTCACAACAACAGCCCTCTGTGCAATTCATG 300  
QY 361 tcaactaactaatttaacaactcttcaagtcctggaattataaacaactcggctatcagga 420  
Db 301 TCACTGACTAATTTTCAACTCTTCAAGTCTGATTTATCAACAAATCGGCTATCGAGA 360  
QY 421 ccaagtcagcaataatgctcaattctcttggtttacacatcaactgaatttgccaataattg 480  
Db 361 CCAGTACCGGATTAATGGCTCATCTTTCTTTTACACCAATAGTTTGCCAAATTAATTTG 420  
QY 481 aatttatgtgaccgtaacttgtagagagccctgcctgatatcccccaatttgccaaca 540  
Db 421 AATTATATGTGACCCGTAAGTGGAGGCCCTGCCGATGCCCAATTTCCGACACA 480  
QY 541 ccttcgcttaccacacatcaacaagtagacagccctcagagacaataatggtccactgagct 600  
Db 481 CTTCCGTTATCCACCAATCAACAGATAGAGCTCCAGGACAAATATGTCACATGAGACT 540  
QY 601 attgctggaggaatactgctgtgtgtgtttacatgcttctgtcgaacctggaatgacttt 660  
Db 541 ATTGCTGGGGAATGACTGCTGTGCTGTACGTTGTGCTGACACCTGCATATGGCAATTT 600  
QY 661 gcatgtgctgcgagaaagaaacccgcagaaacattcttgaatgctccgcctgaaagagac 720  
Db 601 GCATGCTGCGGAGAGAAAGAAACCGGAGAAATTTCTTGATGATGCCAGCTGAAGAGAC 660  
QY 721 ccagaaatgcaacttgctcaactgaagaggtttctctcgagaaatgcaagtgcgaacg 780  
Db 661 CCAGAACTGCACCTTGCTCAACGAAAGAGCTTTCTGCGGAATTCGAAGTCGCAACG 720  
QY 781 gatactttagtaccacttggaagaggtgattgtgaaggtgtataaaggagacgtt 840  
Db 721 GATATCTTTAGTACACTACTGGAAGAGGTGATTTGGTAAGGTATTAAGAGACGCTT 780  
QY 841 gctgagtgtcactcttgtagcaggtttaaaggcttaaaggaaagaaacagacccaggtgtag 900  
Db 781 GCTGATGCTGCACCTGTAGAGATTAAAGCTTAAAGAAAGAAAGAAACACGAGTGGTAG 840  
QY 901 ctgagagttcaacagaggtggaatgattagcatgagctgtgcatcgaactcttctggt 960  
Db 841 CTGAGAGTTTCAACAGAGGTGGAATGATTAGCATAGCTGTGATCATGAAATCTTGTGGCT 900  
QY 961 ctacgtgttcttgcatgacacacagagcggtctctgtatatacatcaatgactaat 1020  
Db 901 CTACGTGTTTTCGATGACACCAACAGACGCGCTCTGTATATCCATATGCTGCTAAT 960  
QY 1021 ggaagtcttcgtcggttttaagagagctgaagcaatagaaccccccctgattgagca 1080  
Db 961 GGAAGTCTTGCTGCTGTTTAAAGAGCGCTCAACCATCAAGACCTCCCTTGATTTGGCCA 1020  
QY 1081 actagaagaagatgactagaagatcgttaaggagcttcttatttgatgacacattgt 1140  
Db 1081 actagaagaagatgactagaagatcgttaaggagcttcttatttgatgacacattgt 1140



Db 1021 ACTAGAGAGAGATGCTAGATCTGCTAGGGGCTTCTTATTTGCATGACCATGT 1080  
Qy 1141 gatcccaagatctccatctgtatgtaaagcgcaaatatattatgtgcaagaattt 1200  
Db 1081 GATCCCAAGATTTCCCTCGATGTAAAGCTGCMAATATATTATGGACGAGAAATTT 1140  
Qy 1201 gaggctgtctaggtgtatcttggttagctagctcattgattacagaatacccatgtt 1260  
Db 1141 GAGCGCTGTGTAGTGTATTTGGGTAGCTAGGCTCATGATTAACAAGTATCCCATGTT 1200  
Qy 1261 acaactctgttaaggggttaacctgtgggtacatagctcccgagtaacctctcgactgaaag 1320  
Db 1201 ACAACTCTGTAAAGGGGTACTTGGGCTACATAGCTCCGAGTACCTCTCGACTGGAAG 1260  
Qy 1321 tcacagaagaacacgaatgtcttctgtatggattatgctcttagagctcattactgga 1380  
Db 1261 TCATCAGAAGAAGACCGATGTCTTTGGTATGGGATTAATGCTCTTAAGCCTACTACTGA 1320  
Qy 1381 cagaagacttctgactcttgcctctgacagatgatagttatgttctgtgacttg 1440  
Db 1321 CAGAGACTTTTGTATCTTGTCTCGCTTGCAGACGATGATGTATGTGTGGATTGG 1380  
Qy 1441 gttaaaagcctttgaaagagaaagttgagatgctgtgactcctgaactgagaac 1500  
Db 1381 GTTAAAGCCTTTGAAAGAGAAAGTGTGAGATGCTGTGCTCATCTGACCTGGAGAAC 1440  
Qy 1501 aattacatctgacacagaagtgtgagcagcttattcaagtagcattactctgtaccagagt 1560  
Db 1441 AATTACATCTGACACAGAAGTGTGAGACCTTATCAATACATTAATCTGTACCCAGGCT 1500  
Qy 1561 tcggcaatctggggcgcttaagatgtcagaggttagtccgaatgttgaaggtgagcctt 1620  
Db 1501 TCGCAGATGGGCGGCTTAAGATGTCAAGAGTACTCCGATGCTTGAAGGTATGGCTTT 1560  
Qy 1621 gcagaaagtgagcagagtgcaaaagtagaagtagcattccatcaagaagcgtagaattagct 1680  
Db 1561 GCAGAAAGTGGAGAGTGGCAAAAGTGAAGTCAATCATCAAGCAGTAGAATTAGCT 1620  
Qy 1681 ccaatctgaactcttgatgatgactctgactcgcagatattctgttctgttgaatta 1740  
Db 1621 CCACATCAACTCTGTGATGATCTTGAAGTCTGACGACATTAATCTGCTTTTGAATTA 1680  
Qy 1741 tctgttccaagataaacaagcatataaattgtgaatgaattatcttctgtctgaa 1800  
Db 1681 TCTGTCCAAATTAACAGCATATAAATGATGAATTAATTTTATGATTAA 1740  
Qy 1801 aaaaaaaaaaaaaa 1815  
Db 1741 AAAAAAAAAAAAAA 1755

RESULT 3  
LOCUS A67827 2089 bp DNA PAT 05-MAY-1999  
DEFINITION Sequence 32 from Patent WO9743427.  
ACCESSION A67827  
VERSION A67827.1 GI:4756649  
KEYWORDS  
SOURCE  
ORANISM  
thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsi.  
1 (bases 1 to 2089)  
De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.  
AUTHORS  
TITLE PRODUCTION OF AROMATIC SEED  
JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
Location/Qualifiers  
1..2089  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"

CDS  
/clone="SERK GENE CDNA"  
195..2072  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB4254.1"  
/db\_xref="GI:4756650"  
/translation="MESSYVFILSLILPNHSIMLASANIEGDALHTLAVTLVDPN  
NVLQSDPPLVNDCTWFHVTCCNNNSVIRVDGNALSGHLVPELGLVNLQLELYS  
NNITGPISNLGNLNLVSLDYLNSFSGPIPEISLGLSLKSLRLNLNNSLGLSIPMS  
LNITTLQVLDLSNNRLSGSVDPNGSLFEPISFANNLIDLCGPVTSHPDGPSPFP  
PPPIOPPPVSTPSPGYSIGTGAIVGAAALPFAAPAIAPAMRRSRSLPIFDVPA  
EBDPEYHLQIKRSLRELOVASDGSFNKNTLRGFGKTKVYKGRADGTAAVRRIKE  
ERTPGSELQFQTEVEHISMAVRYNRLRLRGFEMTPELILVYPYANSVASCERP  
PSPPLDMPTRKRIALGSARGSLYLHRCDEPILIRDKAANILDEEFVAVGDFGL  
AKLMDYKDHVTVAVNGTIGLTAPELSTGKSSKEDVFGIMLELITGQAFDLA  
RLANDVMLLDVWKLLEKRLKMLVDPDQTNVEERLEQVILVALCTQSGPMER  
PKMSEYVRMLEGDGLAEKMDQKVEILREEIDLSLPNPSMDILSTYNLHAVEISGP  
R"

BASE COUNT 569 a 419 c 502 g 599 t  
ORIGIN

Query Match 51.5%; Score 934.8; DB 5; Length 2089;  
Best Local Similarity 74.5%; Pred. No. 4,5e-199;  
Matches 1192; Conservative 0; Mismatches 402; Indels 6; Gaps 1;

Qy 175 ttggaagcttacagaacaataagaagccaattccctagtgattcttgggaatctgaca 234  
Db 486 TTGGAGCTTTACACTAACATTAACATTAAGCCGCAATCTCAATATCTTGGAAATCGACA 545  
Qy 235 aatttggtgaagcttggaacctatcatgaatagcttctctggaacttatccagacatca 294  
Db 546 AACTTGAAGATTTTGAATCTTACTTAACAAGCTCTCCGGTCTCTATTCGGAGATCATGG 605  
Qy 295 ggaagagcttaagaagctaatctctgctctcaaaacaagaagcctctctgtgccaat 354  
Db 606 GGAAGCTTTCAAGCTGAGATTCTCCGGCTTAACACACACAGTCTCATCTGGGTGAAT 665  
Qy 355 ccaatgtacactgaactatatacaactctcctaagctccctggaattatcaacaacatcgcta 414  
Db 666 CCTATGTACTGACCAATATTACTTACCTTCMAAGTGTTAATCTATCAATAACACATC 725  
Qy 415 tcaagacagtaacggatgaatgctatcttcttggttacacatcaagatttggcaat 474  
Db 726 TCTGTAGAGTCCCGAATAGCGCTCCTTCACCTTCACTTCAACCCATCAATTTGCTAAT 785  
Qy 475 aattgaattatgtgaccccgtaactcgtggaagcgccgcctggaatctccccaatttg 534  
Db 786 AACTTGAACCTATGTGACCTGTTAACAAGTCACCAATGCTCTGATCTCCCGCTTTCT 845  
Qy 535 ccaccacctcgttcatccacatcaacaagtaacagctccacaggaacaatgtgtccact 594  
Db 846 COTCCACCACTTTTATTAACAACCTCCCGAGTTTCCACCCCGAGTGGATGATTAAT 905  
Qy 595 ggaactatctgtggggagtagctgtcgtgtgcttcttactgttgggtgacatgcaatg 654  
Db 906 GGAGCAATAGCTGTGGGTGCTGCAAGGTCTCTTGGCCCTTGGCTGTGCAATA 965  
Qy 655 gcaattgtcatgtgctggagagaagaacacgcggaagaactctcttgaatgtgcagctgaa 714  
Db 966 GCCTTGTCTGTGTGGCGAGCAAGAAAGCCCACTAATATTTTCTTCGATGTCCCTGCCGA 1025  
Qy 715 gaggaccagaagtgacacttgctgaactgaagagtttctctctggaagatgtcaatg 774  
Db 1026 GAAGATCCAGAGTTCACTGTGGACAGTCAAGAGGTTTCTTCTTGGCGGAGCTACAGTG 1085  
Qy 775 ggaagagtaactttagt-----acatacttggaaagagtggaattgtgtaagtgat 828  
Db 1086 GCGAGTGAATGGTTTGAATAACAAGAACTTTTGGCGAGAGGTGGTTTGGGAAAGTCTAC 1145  
Qy 829 aaggaagccttctgtatgctcaatctgttagcaagttaaagagcttaagaagaacgaca 888  
Db 1146 AAGGACGCTTGACAGAGAACTCTTGTCTGCTCAAGAGACTGAAGAGAAAGCAACT 1205



QY	889	ccagtgctgtaagctgaattcaaaacagagcgagaaattgattagatgctgctgctacga	948
Db	1206	CCAGGTGAGAGGCTCCAGATTCCAAACAGAAAGAGATGATTAAGATGAGGACGTTTCATCGA	1265
QY	949	aattcttcgcttcaagctgttctctgcatgacaccaaagacggctcttgatatacca	1008
Db	1266	AACCTGTGAGATTACGAGGTTCTGTATGACACCGACCGAGATGCTGTGTATCCT	1325
QY	1009	tacatgctaaatgaagctgttcgctctgtgtttaagaagcgctcaagtcagaaaccccc	1068
Db	1326	TACATGGCCAAATGAAGTGTGTGTTGCTTCACAGAGAGGCCCCACCTCCAAACCTCCG	1385
QY	1069	cttgattgccaactaggaagagattgcaactagagatcctgctaaagggtcttcttaattg	1128
Db	1386	CTTGATTGGCCAAACCGGAGAGAAATCCGGCTAAGGCTCAGCTCAGGTTGTCTTACTTA	1445
QY	1129	catgacattgtgatacccagaatatccactgtgatatgaataagctgcgaatatatttg	1188
Db	1446	CATGATACATGCGCATCGAAGATCATCTCCGTGACGTGATAAAGCGCAAAATCCTCTTA	1505
QY	1189	gacgaagaattcagagctgtgtctaggtgaatttggttagctagctcatgattacaag	1248
Db	1506	GAGCAAAATTCGAAGGGGTTGTGTGAGATTCCGGTTGGCAAACTTATGACTATATAA	1565
QY	1249	gatacccaatttacaactctgtaagggttaacctgggctacatagctccagatactc	1308
Db	1566	GACACTACGTGACAAACGACAGCTCGTGGCACCATTCGCAATGCGCCAGAAATATCTC	1625
QY	1309	tcgacttgaagaatcatcagaagaacccgattgcttgtgtatgagattatgctttagag	1368
Db	1626	TCAACGGGAAATCTTCAGAGAAACGACGATTTTCGATACGGAATCATGCTCTTAGAA	1665
QY	1369	ctcattactggaacagagagcttltgattctgtctcgcttcgacagatgatgtatgtatg	1428
Db	1686	CTAATCCACAGACAAAGAGCTTTCGATTCGCTCGGCTAGCTAACGACGACGAGCTATG	1745
QY	1429	tgtgtgatttgggttaaaagccttttgaagaagaaatttgagatgctgtgcatact	1488
Db	1746	TTATTTGACATGGGTGAAGAGATTTGTTGAAGAGAAAGAACTTGAATTTTATGTGATCCA	1805
QY	1489	gaacctggaagaacaaattacattgacacagaagtgtgacagctatctaaatgatactc	1548
Db	1806	GATTTTAAACAACATCGAGGAGAGACATGGAACAAGTATTCAAAGTGGCGTTGCTA	1865
QY	1549	tgtaaccagggtttgcgaatggagcgcgcttaagatgtcgaagtgtatgtcgaatcctgaa	1608
Db	1866	TGCAACGCAAGATCACCAATGAGAAACCAAAAGATCTCGAAGTTGTGAAGATGCTGCA	1925
QY	1609	ggctagtgcccttgcagaagaagtggagacgagtggcaaaaagtatgaagatcatcataagac	1668
Db	1926	GGAATATGGGCTTGCGGAGAAATGGGACGAAATGCAAAAAGTTGAGATTTTGAAGAGAG	1985
QY	1669	gtagaattagctcacaactgcaacttctgaatgatcctatgactgacagataacttgcatt	1728
Db	1986	ATTGATTTGAGTCTTAATCTTAATCTGATTTGATTTCTTGATTTCTTACATTACATTTGCA	2045
QY	1729	gcttttgaattatctgltccaagaataaacaagcatataaa	1768
Db	2046	GCCGTTGAGTATCTGTGCTCCAAAGATTAATAAAAAAAAAA	2085
RESULT 4			
LOCUS	A67796	6695 bp	DNA
DEFINITION	Sequence 1 from Patent WO9743427.		
ACCESSION	A67796		PAT
VERSION	A67796.1	GI:4756622	
KEYWORDS	.		
SOURCE	carot.		
ORGANISM	Daucus carota		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumphylophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		

REFERENCE	1 (bases 1 to 6695)	editors: Asteridae; eumasterids II; Apiales; Apilaceae; Daucus.
AUTHORS	De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.	
TITLE	PRODUCTION OF APOMITIC SEED	
JOURNAL	Patent: WO 9743427-A 20-NOV-1997;	
	CIBA GEIGY AG (CH)	
FEATURES	Location/Qualifiers	
source	1..6695	
	/organism="Daucus carota"	
	/db_xref="taxon:4039"	
BASE COUNT	1845 a 1182 c 1243 g 2421 t	4 others
ORIGIN		
Query Match	30.9%; Score 560.2; DB 5;	Length 6695;
Best Local Similarity	81.9%; Pred. No. 2.7e-115;	
Matches 735; Conservative	0; Mismatches 18; Indels 144; Gaps	2;
QY 1044	agaagcgtcacaccatcagaaacccctcccttgatctggcccaactagaaagagattgcactagg	1103
DB 5800	AGAGCGTCAGCCATCAGAAACCTCCCGATGATGGCCAACTAGGGAAGAGATTCGACTAGG	5859
QY 1104	atctgcctaaaggggcttcttatttcacatgacccattgcatctcccaagattatccatgtga	1163
DB 5860	ATCTCTAGGGGGCTTCCTCAATTCATACACCATTTGATTCCTCAAGATTTCCATCGCA	5919
QY 1164	tgtaaaagctgcgaatataatttttgacagaaagattgaagctgttgaagtattttg	1223
DB 5920	TGTTAAAGCTGCAGAAATATTTATTGGACGAAGAAATTGAGCTTTGTAGTGATTTTGG	5979
QY 1224	gttagctaggctcactgattacaaagatacccatgctttaaactctgttaaagggtaacct	1283
DB 5980	GTTAGCTAGGCGCTCATGATGATTACAGAGATACCATTTCACCTCTGTAAGGGGTACAT	6039
QY 1284	gggctacatagctcccgagatccctctgacctgggaaatccatcaagaagccgatttctt	1343
DB 6040	TGGGACATATGCTCCGAGTACCTCCGACTGGAAAGTCATCAGAGAAAGCCGATGCTT	6099
QY 1344	tggttatgggattatctctcttagagctcattacttgacaagagacttttcatcttgctcg	1403
DB 6100	TGTTATGGGATTAATGCTCCTAGAGGTCAATTCAGACAGAGGGCTTTGATCTTGCTG	6159
QY 1404	ccctggagaaagatgatatgtatgtttgttgatgg-----	1440
DB 6160	CTTTGCGAAGAGATGATGATTTATGTTGTGATGGGTATGTGCCGGGTGTCCTTT	6219
QY 1440	-----	1440
DB 6220	GGTTAATTATTCACATATTAGTGCTTACTACTTGTGTGTGCGCCCTTGTTTATTTC	6279
QY 1440	-----g	1440
DB 6280	TGCGTATTATTGATTTCTTATGTCATGTTATGCAATTTGACCTTTCAAATGCTTTTAG	6339
QY 1441	gtttaaagcccttttgaagaagaaaagttggagatgcgtgcagatccttgacctggagaac	1500
DB 6340	GTTAAAGCCTTTTGAAGAGAAAAGTTGGAGATCTGTGTGATCCCTGCACCTCCAGAAC	6399
QY 1501	aattacattacacagaagttgagcagcttacttaagaagtacattactctgtaccagaagt	1560
DB 6400	AATTACATTGACACAGAAAGTTGAGGCTTATTCAAGTAGCATTTACGTGTACCCAGAGT	6459
QY 1551	tcggcaattgagagggccttaagatgcttagagaggtatgtccgaatgcttgaagtgatggcctt	1620
DB 6460	TGCGCAATTGAGAGGGGCTTAAGATGTGAGGTGATGTCGCAAAAGCTTTGAAGGATGGCTT	6519
QY 1621	ggcaaaagctggagcagatgtgcaaaaagtagagatcattcaatcaagaacgttagaattagct	1680
DB 6520	GCAAAAAGTGGGACAGAGTGGCAAAAAGTTGAATGATCCATTCAGAGAGTATGAAATTTGCT	6579
QY 1681	ccacatcgaaactctgaaatgatccttagactcgacagataaacttgcatgctttgaatta	1740
DB 6580	CCACATCGAATCTTCTAAATGATCTTCAGACTCGACAGATTAATCTGATGCTTTGATTTA	6639



1741 tctgtccaaagataaacagcatataatgtgaaatgaataatttttaagtt 1797  
|||||  
Db 6640 TCTGTCACAGATTAACAGCATATTAATGT-ATGCAATTAATTTTATGCTT 6695

RESULT 5  
F23M19/c  
LOCUS  
DEFINITION Arabidopsis thaliana chromosome 1 BAC F23M19 sequence, complete  
ACCESSION AC007454  
VERSION AC007454.3 GI:5091613  
KEYWORDS HTG.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;  
Arabidopsis.  
1 (bases 1 to 88401)  
REFERENCE  
AUTHORS Vysotskaya,S.,Schwartz,J.R.,Yu,G.,Toriumi,M.,Lenz,C.,Liu,S.,  
Lee,J.,Liu,A.,Li,J.,Kremenetskaya,I.,Luros,J.,Gonzalez,A.,  
Altafi,H.,Araujo,R.,Chao,Q.,Conn,L.,Conway,A.B.,Dunn,P.,  
Hansen,N.,Hulzar,L.,Kim,C.,Palm,C.,Rowley,D.,Shinn,P.,  
Walker,M.,Davis,R.W.,Ecker,J.R.,Federspiel,N.A. and Theologis,A.  
Arabidopsis thaliana chromosome 1 BAC F23M19 sequence  
Unpublished (1999)  
2 (bases 1 to 88401)  
REFERENCE  
AUTHORS Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
3 (bases 1 to 88401)  
REFERENCE  
AUTHORS Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUN-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
4 (bases 1 to 88401)  
REFERENCE  
AUTHORS Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUN-1999) Plant Gene Expression Center, 800 Buchanan  
St., Albany, CA 94710, USA  
On Jun 17, 1999 this sequence version replaced gi:4996902.  
The sequence of BAC F23M19 from Arabidopsis thaliana chromosome 1.  
FEATURES  
SOURCE  
1. 88401  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F23M19"  
/complement(796..3891)  
/gene="F23M19.1"  
complement(join(796..1116,1205..1271,1349..1432,  
1506..1636,1794..1850,1999..2047,2166..2251,2468..2545,  
2771..2893,2990..3079,3154..3254,3332..3393,3467..3556,  
3638..3891))  
/note="F23M19.1"  
/note="Contains a pfl00561 alpha/beta hydrolase fold  
domain."  
/evidence-not\_experimental  
/protein\_id="AAD39602.1"  
/db\_xref="GI:5091614"  
/translation="MEEDILSPDLDLFOALTLPIRHYLALFVLTITFNV  
EFHLDGALIRFCRNLINPDSPIHGVSKCOLHGRVATPWLPASPHQTEP  
NFHGLPFTYTRQLFLTSDGTLALDMLNSDLVSNKSEITKEDTPIAVVP  
GLTSDSSAYLKLHAYDAKTGMNVVSNRGLGVSATSDCEFNAGMTDIDIVADY  
LQHKYRAPLEIGTSIGANLYKVLGEEKPLRGAVLAICSPMDLIDRFLCRTL  
KOKLYKRAITIGOGYAOIHQEPFLRANWGIKRSIRDPDNHATCIYKFEETAT  
YKRSSTQYGVNVAVPLICTSALDDPLCTREARIPDECRANKKIYATATNNGHGLAF  
FEGLTGSLMWVRAITNEFLGLVSCSPYMHIOKIVDKRSSGSGKOEPIINOGPYLTIAE

gene  
CDS  
complement(5693..6732)  
/gene="F23M19.2"  
complement(join(5693..6108,6233..6428,6499..6732))  
/note="Similar to gb|X98322 peroxidase (pxrx10) from  
Arabidopsis thaliana."  
/codon\_start=1  
/evidence-not\_experimental  
/protein\_id="AAD39603.1"  
/db\_xref="GI:5091615"  
/translation="MRAITASFIFLCPVPSYRGKLRGEFYDSSCPLEASTVAVYAK  
HSLNQVYTAALRMQFHCDFVNYRKILLCFVYGCDSALIDSTEPEPSKSAEAY  
SVRGEIIDEVKLEIYCPPEVSCADIVTATFDSIALAGLVNTMALIGGSHV  
GVACISLQDRLKDPKMDSKLAKLQTCGNDPDSVLDQMTPLEVDNIOYKQISO  
GIALRIDNGLDLDSTSRIVSNFALNTEITGERBAEMQMGELKYLVTGNSGEIRNC  
IVVSH"  
complement(10781..14114)  
/gene="F23M19.3"  
complement(join(10781..11318,11416..11460,11670..11746,  
11875..11991,12207..12310,12443..12512,12614..12717,  
12848..12941,13174..13213,13394..13719,13800..14114))  
/gene="F23M19.3"  
/codon\_start=1  
/evidence-not\_experimental  
/product="F23M19.3"  
/protein\_id="AAD39604.1"  
/db\_xref="GI:5091616"  
/translation="MGLGCSRSSVNNAPGCTAHVNGHLLNNASDLNHSGBSGIK  
DDPSPEVENVDDNKRHTSESFSFPIVSSGSHSPQNTEDIPRLISRLSOKSSTSRQAA  
VAKVSEVSSLGRAGTGLGKAVDVLDSHSMNLMISGFSATTVKNTKLSISF  
EVANTIVGANIMHLSKDSITLHKEYVLESEGVQNLISQMDIELILAAADRFFQD  
PVTLVYTORDLYHEHLADREDDYORKEEENPSTAOGVDTLAILTELSOKK  
HYRNLRKKSLSRLIEEYMEKLVNVVPHRIHEHEAFGADPDKPADPDINKKKEIT  
AGLALHAKITITQIDLTYSRSTPASTRDLAYGDEPSTKSALRSKISGFYKEELT  
VPQIAEMERKTLQMLVPVAVNTTAHAGFGVGVGMASSSEANQRPAGOTILRDTLH  
HADKEKTEATYILDVYVMLHLVTVQVRAITGGLSPKSPISRNQKOTIOLSSGSHNP  
SMGLPLTTEDEQEMLRDYSKRRTPGISQSEFTVAKARLCKHHRLSKSSHPMAG  
EMKKKNDTSFRRSSVPIIDFDIRKALVIDRDTIRSL"  
23307..26279  
/gene="F23M19.4"  
complement(join(23307..23355,23431..23627,23656..23751,23850..23906,  
24002..24145,24225..24309,24401..24491,24571..24735,  
24842..24955,25045..25117,25210..25348,25453..25865,  
25953..26140,26231..26279))  
/gene="F23M19.4"  
/note="Similar to gb|AF082176 auxin response factor 9 from  
Arabidopsis thaliana."  
/codon\_start=1  
/evidence-not\_experimental  
/protein\_id="AAD39615.1"  
/db\_xref="GI:5091627"  
/translation="MESGVYVNAQPELSGITDGSKSYVEQLMKLCAGPLDIPKLG  
KYTFPQGHIEVLSLSLPLPFSLSLHLSLSLSVETSTEELENEQDIPDLAS  
KLQCVIAIHILKVENNSDETYAEITLMDTQVYIPQENQEPPLVNSRTKVLTAAS  
TSAHGFEVPPKKHAIIECLPSLDSMQPLPAQCLAILDIGHNQRNNHYRGFORHLAT  
TGMNAFTTSKRLVAGDIVFVRGEGELRGVIRARRHQGNIPISSIVSICMRGVYA  
SAKHAFDNOCMEYTVYKPRSSKPIYSYKPIEDAYANNKPNPSSRTFMTEGDFEERC  
EGTITGVSDPEGHKCSFMRSLVQWDEFTFPKPKVSPMDRTHMLPAINVPSPFL  
KNKRLREVNIGSSSHLPLPLTLOGQENBLSVASPNNISLRTRDTEDMANPSKLL  
MSYPVQPMKLYNNQWTEMEENTTKTGTFNRLFGVTLDTPVILDPLEIGSEIS  
KLTEGRKFGLSQTLRSPTEIONKQFSSSRCTGKLVQMGVTVIGRAVDLSVNGDQLLL  
ELEKLFIDIKGQJOTRNQWEIAFTDSEDKMLVGDPPWPEFCNMVKKIFIQRR"  
complement(28888..31377)  
/gene="F23M19.5"  
complement(28888..31377)  
/note="Contains similarity to gi479356 protein kinase PK1  
from Zea mays, is a member of the pfl00954 S-locus  
glycoprotein family and contains a pfl00069 Eukaryotic  
protein kinase domain."  
/codon\_start=1



Query Match	23.9%	Score 434	DB 8	Length 88401
Best Local Similarity	64.2%	Pred. No. 4,36-87		
Matches 828	Conservative 0	Mismatches 245	Indels 217	Gaps 3
697	tttgaatgacagctaaagagcccaagaatgacaccttggttaacatgaaagaggtttct	756		
Db 73350	tttcgtgatctacaccccaahgahcctctagcttcttggggcagactttaaagcggtttct	73291		
757	ctcgagaaattgcaagtcgcaacgagatctttag-----taccactactggaagagt	810		

Dd	73230	CTAGGGAACTTCAGTACACACAGTATAGCTTCAGCAACAAGAAACAATT77GGGCGGAGGT	73231
Qy	811	ggatttgtaagtgataaaggagccctgcyaagctcacttgacgaagttaaag	870
Dd	73230	GGGTTCCGAAAGTCTACAAAGCGCGTCTTCATGGAACACTTGTGTGAGCAACGCG	73171
Qy	871	cttaagaagaagacaccagtgtagtgactgaatttaacaagaaggtggaaatgatt	930
Dd	73170	CTTAAGAAAGAGGAAACCCCAAGTGGCAGCTCCAGTTTCAGACAGAGTGGAGATGATA	73111
Qy	931	agcatggtctgcatcgaatctctgcgtacgtyggttctgcatgacacacaagag	990
Dd	73110	AGCATGCGCGTTACAGAAATCTCCTCAGCGCTACGCGGTTTCGTATGACCCCTACGAG	73051
Qy	991	cgagctcttgataatccaacttgcttaagagtggtgcgcgtggtttaag-----	1044
Dd	73050	AGATGGCTTGTATTCTTACATGGCTAATGGAAGTGTCCGTCCTCTTGGAAGGTAAAC	72991
Qy	1044	-----	1044
Dd	72990	CTTGGAATTTTAAGTGTTCATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	72931
Qy	1044	-----	1044
Dd	72930	TTGTTTGTATCTTAACCAATTTTAAACCTTCCAAATACAAATGAGTACCTTCTGAA	72871
Qy	1044	-----	1095
Dd	72870	ATGATTACAAAGCTCCACCATCACAGTTGCCCTCTAGCCTGGTCAATMAACAGCAATTC	72811
Qy	1096	gcaataggaatctgctagggggtcttctatttgatgacagccatgtgatlccagaattac	1155
Dd	72810	GCCTAAGATACACGAGGGGTTGCTTATCTTCATGATCATTCGACCCCAAAATTAAT	72751
Qy	1156	catgtgatgtaaagctgcaatatattatggacgaagaattgggctgtgtgaggt	1215
Dd	72750	CACCGTATGTGAAGTGTCTAATATCTGTGTGACGAGGAAATTTGAGGCGGTGATAGGT	72691
Qy	1216	gatttgggttaagctaggtcctatgatacaagaataccatgttataacgtctgttaag	1275
Dd	72690	GATTTCGGGTAGCTAGACTTAAGACTTAAGACTTAAGACTTAAGACTTAAGACTTAAG	72631
Qy	1276	ggtaccttgggtctacatagctcccgagtaactctctgactgtyaaagtcataagaagacc	1335
Dd	72630	GGGACTTTTGAACAACATTGCTCTGTGATCTCTCACTCACTGGAATAATCTTCAAGAAATCT	72571
Qy	1336	gaigtcttggtatggatgatctgctcttagagctcatctacttgacagagagcttgat	1395
Dd	72570	GATGTTTTTGGCTACGGGATCAAGCTTTTGGAACTGATTAAAGGTCGAGAAGCTTTTGAT	72511
Qy	1396	cttgctcgcccttgcgaacgaatgatagtatgtatgttgttgaatg-----	1440
Dd	72510	CTTGCAAGACTGCGGATGACGATGACGATTATCTCTCAAGATTGGGTATTAACACAGATCT	72451
Qy	1440	-----	1440
Dd	72450	TTTAACACATATCTGCTATCTCTCAAAAAGCTGATTAATCTGTTCAATTTGGTCTTCTCA	72391
Qy	1440	ggttaaaagcccttgaagaagaagaagttgtagatgtctgtgcactctgaacctgtagaag	1499
Dd	72390	GGTGAAGGCTTTTGAAGGAAAGAAAGCTGGAGATGCTTGGGATCCTGTGACCTGCAAAAG	72331
Qy	1500	caattacattgacacagaagttagagcagctatccaagtaagcatctctgtlaccaggg	1559
Dd	72330	CATTTACACAAACAAACAAATTAACAAAGCTCAATTAAGTGTCTTCTGCAACACAGAG	72271
Qy	1560	ttcgcaatggagcggtcctaagaatgcacaaggtagtcgaaatgcttgaagttagagcc	1619
Dd	72270	CTCACCTATGGAAGAGACTTAAGATGCTCTAGGTTGTTCGATTCGTTGAAGTACGGTTT	72211
Qy	1620	tgacagaaagtggagcagatgacaaaagttagaagtcacatcaacaagacgtagaatgac	1679
Dd	72210	AGCGGAGAAATGGGACACAGTGGCAGAAAGTGGAAAGTCTTCAGGCAAGAGTGGAGCTTTC	72151











Oy	1648	gtagagatcatccatcaagacgtagaatctgcacc	c	1687
Db	50080	GAGGAAATGTTTCAGACAGAATTCTTCAACTACCACCCACC	C	50041
RESULT	7			
LOCUS	AC012654	98471 bp	DNA	PLN
DEFINITION	Arabidopsis thaliana chromosome 1 BAC F14O23 sequence, complete			10-DEC-1999
VERSION	AC012654			
KEYWORDS	HTG.			
SOURCE	thal. cress.			
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
REFERENCE	1 (bases 1 to 98471)			
AUTHORS	Liu,S.X., Lee,J.M., Sakano,H., Yu,G., Jhaeri,A., Lenz,C., Toriumi,M., Chin,C., Chlou,J., Choi,E., Gonzalez,A., Hwang,B., Koo,T., Li,J., Liu,A., Pham,P., Vaysberg,M., Altafi,H., Brooks,S., Buehler,E., Chao,Q., Conn,L., Conway,A.B., Hansen,N., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Nguyen,M., Palm,C., Shinn,P., Tambunga,G., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A. Arabidopsis thaliana chromosome 1 BAC F14O23 sequence			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 98471)			
TITLE	Theologis,A.			
JOURNAL	Direct Submission			
AUTHORS	Submitted. (02-NOV-1999) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA			
REFERENCE	3 (bases 1 to 98471)			
TITLE	Theologis,A.			
JOURNAL	Direct Submission			
AUTHORS	Submitted (10-DEC-1999) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA			
JOURNAL	On Dec 10, 1999 this sequence version replaced gi:6175131.			
COMMENT				
FEATURES	Location/Qualifiers			
source	1..98471			
	/organism="Arabidopsis thaliana"			
	/cultivar="Columbia"			
	/db_xref="taxon:3702"			
	/chromosome="1"			
	/clone="F14O23"			
BASE COUNT	31501 a 17757 c 17475 g 31738 t			
ORIGIN				
Query Match	21.8%; Score 395.6; DB 8; Length 98471;			
Best local Similarity	62.6%; Pred. No. 1.6e-78;			
Matches 787; Conservative	0; Mismatches 274; Indels 197; Gaps 3			
Oy	707	cagctgaagaaggaccagaagtgcacctggttcaactgaagagtttctctcgsgaat	766	
Db	79858	CACCACCAAGAAAGATCCAGAAAGTTCATCTCGGACAGCTCAAGAGGTTTTCTTGCGGGAGC	79917	
Oy	767	tgcagtcgcaaaaggatacttaagt-----acatacttggaagagtgatattgta	820	
Db	79918	TACAAGTGGCGAATGATGGTTTAGTAACAAGAACNTTTTGGGCAGAGGTGGGTTGGGA	79977	
Oy	821	aagctgataaaggacgccttgtcgtatgctcacctgtgatagcagttaaaggcttaaagaag	880	
Db	79978	AAGTCTCAAGAGGACGCTTGCGACAGGAACTCTTGCTGCTCAAGAGACTCAAGGAAG	80037	
Oy	881	aagcaaaccaaggtgtgtagctgcagtttcaaacacagagtgagaaatatgagatgctg	940	
Db	80038	AGCGAACTCCAAGGTGGAGACTCCAGTTTAAACAGAAAGTAGAGATCATAGTARITGGAG	80097	
Oy	941	tgcatacgaatatctctcgtctacgtgtgttcttgatgataccaacaagacggactcttg	1000	

Db	80098	TTCAATGAAACCGTGGAGATTAGAGGTTTCGTATGACACCGACGAGAAATTGCTTG	80157
QY	1001	tataccatacacygctaagtgaagctgttcgctcgcttaag-----	1044
Db	80158	TGTAATCTTACATGAGCGCATATGAAATGTTGCTTCCTGTCTCAAGAGGTAATAACTAAACAT	80217
QY	1044	-----	1044
Db	80218	CTTGCTCTCTCTCAATTACTTTCAGCTGAAGCTTTTTCATGTTTCCTTTANGGT	80277
QY	1044	-----	1075
Db	80278	TCATAAATTGGTGTACATAATGACACAGAGAGGCCACCGTCACAAACCTCCGTTGATT	80337
QY	1076	ggccaactatggaagagatgtgactaggaatctgttaggggctcttctatcttgcatgacc	1135
Db	80338	GGCCAAAGCGGGAAGAAATCGCGCTAGGCTAGCTCAGAGTTGTCTTACTCAATGATC	80397
QY	1136	attgtatcccaagatcatccatcgtgtatgtaaaagctgtcaatatatatttgcagaaag	1195
Db	80398	ACTCGAATCGGAAGATCATCTACCGCTGACGTAAAGACGAACAACATCCTTTAGACGAAG	80457
QY	1196	aatttgagctgtgttaggtgatttttgggttaagcttaggtccatgagattaaagatacc	1255
Db	80458	AATTGGAAGCGGTTGTGGAGATTTCGGGTTGGCAAAAGCTAAATGAGCTATAAAGACATC	80517
QY	1256	atgtcaactctgtctgaagggttacctctgggtcacaatagctcccgagtaacctcgactg	1315
Db	80518	ACGTGACAAACGACATGTCGCTGGCCACCATCGGTGCACATGCGTCCAGAAATATCTCAACG	80577
QY	1316	gaagatcatcagaagaaccgatgctcttggattatgagatagatgaccttagagctcata	1375
Db	80578	GAAATCTTCAGAGAAACCGACGATTTTCGGATACGGAAATCATGCTTCTGAACATAATCA	80637
QY	1376	ctgagagagagacttttgatctcttgcgccttgcgaagcatgatatgattatg-----	1429
Db	80638	CAGGACAAAGCTTTCGATCTCGCTCGGCTAGCTCAACGAGACGACGATGATTTACTTG	80697
QY	1429	-----	1429
Db	80698	ACTGGTAAATACACATAAACTTCATTTGACCCCAACCAAAACCATTAAGAAACAGAA	80757
QY	1429	-----	1469
Db	80758	TCTAATCTCAAAATCTTGTGTTGTGTTTAGTGAAGAGATTGTTGAAGAGAAAGACT	80817
QY	1470	ggaagatgctgctcatctccatcccttgcgagacaattatcatctgacacagaagttgagcagct	1529
Db	80818	AGAGATGTTAGTGATCATCATCTTCAACCAAACTACAGAGAGAGAACTGGAACAAAT	80877
QY	1530	tattcaagtagcatatctctgtatcccaaggtctgcgaatgtgagagggcttaagtgtcaga	1589
Db	80878	GATTAACAATGGCGTTGCTATAGCCAGCAAGAAATCACCAATGGAAGACCAAAAGATGCTGA	80937
QY	1590	ggttagctcgaatgctttaaagtgatgctcttgcagaagaagtggagcagtgctgcaaaagt	1649
Db	80938	AGTTTATAGATGCTGGAAGAGATAGGCTTTCGGAGAAATGGACGAATGGCAAAAAG	80997
QY	1650	agaagtcacatcaagaagctagaattagctccacatcgaactctctaatgagatcctaga	1709
Db	80998	TGAGATTTTGGAGGAAAGATGATTGTTGATCTCAATCTCACTGATTAATGGATTCTGA	81057
QY	1710	ctcgacagatgactgcatgcttcttgaattatctgtgtccagagataaacagcatataa	1767
Db	81058	TTTCACTTACATTGGACCGCGTTGATGTTATCTGTGTCACAGATRAAAAAAAAAAACHA	81115



ACCESSION AC006436 AE002093  
 KEYWORDS AC006436.4 GI:6596565  
 SOURCE HRC.  
 ORGANISM *thale cress*.  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 eudicotyledons; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsids.  
 REFERENCE 1 (bases 1 to 65899)  
 AUTHORS Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D.,  
 Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E.,  
 Feldhym, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Ronning, C.M.,  
 Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., VanAken, S.E., Umayam, L.,  
 Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H.,  
 Goodman, H.M., Somerville, C.R., Cohenbayer, G.P., Preuss, D.,  
 Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and  
 Venter, J.C.  
 TITLE Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana  
 JOURNAL Nature 402 (6763), 761-768 (1999)  
 MEDLINE 20083487  
 REFERENCE 2 (bases 1 to 65899)  
 AUTHORS Lin, X.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-DEC-1999) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA  
 COMMENT On Dec 17, 1999 this sequence version replaced gi:4726109.  
 The sequence and annotation of chromosome 2 were merged from those  
 of the individual clones on this chromosome after removing  
 overlaps. For detailed information, please see the TIGR web site  
 (<http://www.tigr.org/tdb/at.html>).

Genes were identified by a combination of three methods: Gene  
 prediction programs including GRAIL,  
 (<http://archur.epm.ornl.gov/pub/xyrall>), GeneFinder (Phil Green,  
 University of Washington), Genscan (Chris Burge,  
<http://genomic.stanford.edu/GENSCANW.html>), and NetPlantGene  
 (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the  
 complete sequence against a peptide database and plant EST  
 databases at TIGR, and manual curations based on those analyses.  
 Annotated genes are named to indicate the level of evidence for  
 their annotation. Genes with similarity to other proteins are named  
 after the database hits. Genes without significant peptide  
 similarity but with EST similarity are named as 'unknown' proteins.  
 Genes without protein or EST similarity, that are predicted by two  
 or more gene prediction programs over most of their length are  
 annotated as 'hypothetical' proteins. Genes encoding TRNAS are  
 predicted by tRNAscan-SE (Sean Eddy,  
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were  
 identified by RepeatMasker (Arian Smil,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are  
 numbered from the top to the bottom of the chromosome.

We thank the GSH/Mashu/ABI consortium for sequencing BAC clones  
 F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone  
 F13J4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards  
 and Satoshi Tabata for helpful assistance. In addition, we would  
 like to thank the TIGR Bioinformatics Department, especially Lixin  
 Zhou, Hanif Khalek, Michael E. Heeney, Lily Fu, Feng Liang, Jeremy  
 Peterson, Michael Holmes, and Delwood Richardson for software and  
 database support.

This work was supported by the National Science Foundation,  
 Department of Energy and the US Department of Agriculture.

Address all correspondence to: [at@tigr.org](mailto:at@tigr.org).

FEATURES  
 source  
 1.65899  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="2"

misc\_feature  
 1..65698  
 /note="Sequence from clone F13J11"  
 /complement(join(451..842,907..1234))  
 /gene="At2g13660"  
 gene  
 complement(451..1234)  
 /gene="At2g13660"  
 /note="F13J11.1"  
 CDS  
 complement(join(517..842,907..1114))  
 /gene="At2g13660"  
 /note="unknown protein"  
 /codon\_start=1  
 /protein\_id="AAD28310.1"  
 /db\_xref="GI:4726110"  
 /translation="MOCGMREKGFATVDSSSPFSICRLITGNFLHQPADOL  
 SPNNYLICRLVYLARCHSPETIODENSSSVPIPNRLINLPRTALANKPF  
 WRSELSISGRIRIKHSNNKDPDPFNHFLPEIGRVNLSLQRAIKRKKNVKNSEK  
 ETGIKGNLHKPKSSD"  
 complement(2606..2687)  
 /rpt\_family="GAA)n"  
 join(<3349..3541,3644..3744,4109..4230,4375..4471,  
 4583..4693,4792..4928,5629..5755,5848..5997)  
 /gene="At2g13670"  
 <3349..5997  
 /gene="At2g13670"  
 /note="F13J11.2: predicted by genscan and genefinder"  
 join(3349..3541,3644..3744,4109..4230,4375..4471,  
 4583..4693,4792..4928,5629..5755,5848..5997)  
 /gene="At2g13670"  
 /note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="AAD28311.1"  
 /db\_xref="GI:4726111"  
 /translation="MAOSTSHDGPQGLARRPSSAAATVSIIEVDHEVPSLGTI  
 APLIRVAETIEREPYAVYLCRFAPFAKRLDPSSGSGRGVQFKTLQRLERDAS  
 SLARKVKTGSEVESFYQYVYALDQGOADQADQKAYOTAGVLEFVMAVN  
 KSEVEYAPPEITLAADYQKNEIYPIYITLDSGASQSMQLEEKAAVALGN  
 TRGLNWSGFEQHRKKTGNDLDMKAMGFQASWPCPADIDQKRLTMEGLYLI  
 WGEANVEMFECICYIFHMAVELHGLAGNVSIVYGENIKPSYGGDEAPLAKVIT  
 PIYRVQYVQV"  
 10367..13711  
 /gene="At2g13680"  
 /note="F13J11.3"  
 /pseudo  
 complement(join(<14929..15843,16401..>17114))  
 /gene="At2g13690"  
 complement(<14929..>17114)  
 /gene="At2g13690"  
 /note="F13J11.4: predicted by genscan and genefinder"  
 complement(join(14929..15843,16401..17114))  
 /gene="At2g13690"  
 /note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="AAD28312.1"  
 /db\_xref="GI:4726112"  
 /translation="MGDSDLRKPNFECTGLHSRRWSCSEAFAPASPDPRSNNH  
 IPAKQOQRKLVCSPOSSKSAINYNRDLPRILSPGVSTIDSPYTTMQELET  
 TQEEEDDAVVDSTPNLSESEFRAPKLEVGSGISSEYDARSLKGRNGGVAVLELS  
 LEVLAANDVSGIJAIEKKCSSSSSGLKNCRIEVCVENLGVFREVVELMPEES  
 NVLIKPEYKGVYRAIDVLEVAAGISRAVLSGLKYLEAVPTDEEEKLRLLGIY  
 SFDDAVAEIILARNSETNEIDQSLKTYWSTSDVNPNEKLSVKGLGCKSS  
 VYEKQPIINKEIDYRAGKCCVDSLAKFEQSSSSSKREKRLISSEVNNIML  
 LEIMIDRIAEPEYINGKORLYEMERKSPMYRIVSYTAITFLAKKKRYOCCG  
 EAKAGLEAMFRPMLVDGWLQRCCKKLDKEVEBGGQTLTLTPVKEQYOVFEMFR  
 WFSKGTCEPNLSKAFQIMWRSFLRGVESSSTOR"  
 complement(18082..20371)  
 /gene="At2g13700"  
 /note="F13J11.5"  
 /pseudo  
 23450..29498  
 /rpt\_family="POLY\_A"  
 complement(29932..30854)  
 /gene="At2g13710"  
 /note="F13J11.6"







FEATURES	source
1. 2278	/organism="Sorghum bicolor"
2. 420-425 (1999)	/cultivar="TX 430"
3. 420-425 (1999)	/db_xref="taxon:4558"
4. 420-425 (1999)	/lab_host="X11-Blue"
5. 420-425 (1999)	/tissue_type="leaf"
6. 420-425 (1999)	/clone="cRLK1"
7. 420-425 (1999)	/gene="RLK1"
8. 420-425 (1999)	/note="accumulates in mesophyll cells"
9. 420-425 (1999)	/product="putative protein serine /threonine kinase"
10. 420-425 (1999)	/protein_id="CAB51480.1"
11. 420-425 (1999)	/db_xref="GI:5596996"
12. 420-425 (1999)	/translation="WMRSEITMKLAFVLMILGCGSSSLCAVDSQEVALYEMKQVLY DNRGASDMKNDQSPCYWEVYVNDQNRKVSITLSSGLTGFLSPILAKLTTLQDLK DNNITGTGIPPEFGNTSLTLNLRNNLNLSIPDSLGQISKQILDSHNLGNTPI SEFSNPSPSINDLNIAVNIISGEIPDHLQAHYFNHNLNGOMLFCGEGSTMGTS SKNSKLVIGSIAVYLVYVALVILMOQRARYPELFTIOSONDMLHEFGQIKR FEMRELQIAVNEFSONVLYCGKGRKYKTVLPDPSITAVKRLPNTERRHDEGLATRL REVELSIAVHNKINIRLIGECTPTTELLYPPYEMTSASLRDITAKNEYLDKSTR KTLALGASAGSLYHERHCNPKIIRDYKALNAVLLDSEFNAVYDGLKAMDIGENTV TGVGVNMGHIAPEYIKTGRSPYTDIYGVALLIYVGERALIFHPRDEEGEIM LIDQVYLMWEGRELIDVDNMLGCVNLYLELEFVTOIALICTHMPESQPTMSVEYOM LGEIYPAERMEWEMOIAELORROOHEMRQOKLFPSESLINIOEALETIGSR"
13. 420-425 (1999)	/gene="RLK1"
14. 420-425 (1999)	/gene="RLK1"
15. 420-425 (1999)	/product="putative protein serine/threonine kinase"
16. 420-425 (1999)	/product="putative protein serine/threonine kinase"
17. 420-425 (1999)	/product="putative protein serine/threonine kinase"
18. 420-425 (1999)	/product="putative protein serine/threonine kinase"
19. 420-425 (1999)	/product="putative protein serine/threonine kinase"
20. 420-425 (1999)	/product="putative protein serine/threonine kinase"
21. 420-425 (1999)	/product="putative protein serine/threonine kinase"
22. 420-425 (1999)	/product="putative protein serine/threonine kinase"
23. 420-425 (1999)	/product="putative protein serine/threonine kinase"
24. 420-425 (1999)	/product="putative protein serine/threonine kinase"
25. 420-425 (1999)	/product="putative protein serine/threonine kinase"
26. 420-425 (1999)	/product="putative protein serine/threonine kinase"
27. 420-425 (1999)	/product="putative protein serine/threonine kinase"
28. 420-425 (1999)	/product="putative protein serine/threonine kinase"
29. 420-425 (1999)	/product="putative protein serine/threonine kinase"
30. 420-425 (1999)	/product="putative protein serine/threonine kinase"
31. 420-425 (1999)	/product="putative protein serine/threonine kinase"
32. 420-425 (1999)	/product="putative protein serine/threonine kinase"
33. 420-425 (1999)	/product="putative protein serine/threonine kinase"
34. 420-425 (1999)	/product="putative protein serine/threonine kinase"
35. 420-425 (1999)	/product="putative protein serine/threonine kinase"
36. 420-425 (1999)	/product="putative protein serine/threonine kinase"
37. 420-425 (1999)	/product="putative protein serine/threonine kinase"
38. 420-425 (1999)	/product="putative protein serine/threonine kinase"
39. 420-425 (1999)	/product="putative protein serine/threonine kinase"
40. 420-425 (1999)	/product="putative protein serine/threonine kinase"
41. 420-425 (1999)	/product="putative protein serine/threonine kinase"
42. 420-425 (1999)	/product="putative protein serine/threonine kinase"
43. 420-425 (1999)	/product="putative protein serine/threonine kinase"
44. 420-425 (1999)	/product="putative protein serine/threonine kinase"
45. 420-425 (1999)	/product="putative protein serine/threonine kinase"
46. 420-425 (1999)	/product="putative protein serine/threonine kinase"
47. 420-425 (1999)	/product="putative protein serine/threonine kinase"
48. 420-425 (1999)	/product="putative protein serine/threonine kinase"
49. 420-425 (1999)	/product="putative protein serine/threonine kinase"
50. 420-425 (1999)	/product="putative protein serine/threonine kinase"
51. 420-425 (1999)	/product="putative protein serine/threonine kinase"
52. 420-425 (1999)	/product="putative protein serine/threonine kinase"
53. 420-425 (1999)	/product="putative protein serine/threonine kinase"
54. 420-425 (1999)	/product="putative protein serine/threonine kinase"
55. 420-425 (1999)	/product="putative protein serine/threonine kinase"
56. 420-425 (1999)	/product="putative protein serine/threonine kinase"
57. 420-425 (1999)	/product="putative protein serine/threonine kinase"
58. 420-425 (1999)	/product="putative protein serine/threonine kinase"
59. 420-425 (1999)	/product="putative protein serine/threonine kinase"
60. 420-425 (1999)	/product="putative protein serine/threonine kinase"
61. 420-425 (1999)	/product="putative protein serine/threonine kinase"
62. 420-425 (1999)	/product="putative protein serine/threonine kinase"
63. 420-425 (1999)	/product="putative protein serine/threonine kinase"
64. 420-425 (1999)	/product="putative protein serine/threonine kinase"
65. 420-425 (1999)	/product="putative protein serine/threonine kinase"
66. 420-425 (1999)	/product="putative protein serine/threonine kinase"
67. 420-425 (1999)	/product="putative protein serine/threonine kinase"
68. 420-425 (1999)	/product="putative protein serine/threonine kinase"
69. 420-425 (1999)	/product="putative protein serine/threonine kinase"
70. 420-425 (1999)	/product="putative protein serine/threonine kinase"
71. 420-425 (1999)	/product="putative protein serine/threonine kinase"
72. 420-425 (1999)	/product="putative protein serine/threonine kinase"
73. 420-425 (1999)	/product="putative protein serine/threonine kinase"
74. 420-425 (1999)	/product="putative protein serine/threonine kinase"
75. 420-425 (1999)	/product="putative protein serine/threonine kinase"
76. 420-425 (1999)	/product="putative protein serine/threonine kinase"
77. 420-425 (1999)	/product="putative protein serine/threonine kinase"
78. 420-425 (1999)	/product="putative protein serine/threonine kinase"
79. 420-425 (1999)	/product="putative protein serine/threonine kinase"
80. 420-425 (1999)	/product="putative protein serine/threonine kinase"
81. 420-425 (1999)	/product="putative protein serine/threonine kinase"
82. 420-425 (1999)	/product="putative protein serine/threonine kinase"
83. 420-425 (1999)	/product="putative protein serine/threonine kinase"
84. 420-425 (1999)	/product="putative protein serine/threonine kinase"
85. 420-425 (1999)	/product="putative protein serine/threonine kinase"
86. 420-425 (1999)	/product="putative protein serine/threonine kinase"
87. 420-425 (1999)	/product="putative protein serine/threonine kinase"
88. 420-425 (1999)	/product="put

Db	711	TCCGCAACATCTACTTCAGGCGGCTCATTTACAGGCAACCATTTGAATTGTG	770
Qy	552	cccaaccatcaacagt-----acagctcccaagaacatgctccca	592
Db	771	GCCAAATTTATTTTCATGTGAAGGAGTAGACATATACAGTGGATCGAAAAATTC	830
Qy	553	ctgagcatatgctbgtgggagtagtgcgtcgtgctgcttactgtttgcgtcacctycaa	652
Db	831	AGCTAAAGGTGGTTATGTGAAGCATCTGGAGGAGTAGTACTTTGTGTACTGTACTC	890
Qy	653	tgagcatgtgatgtgctggcgagaga---gaaacccgcgagaaacttcttgatgtgcag	709
Db	891	TAGTTTTCATGTGTGCGCAAGATACGTTATCGGCTGTAATCTTCACTTACGTGTACG	950
Qy	710	ctgaagagagaccacgaagtgcaccttgctgaacatgaagagtttctctgcgagaatgac	769
Db	951	GTCAGATGATCATATATGCTGAGTTTGGCCAAATAAAGCGGTTCTCATGCGAAGCTTC	1010
Qy	770	aagtcgaacaggtactttagt-----acatctctgggaaggtgattgttgtagt	823
Db	1011	AGATTGCAACCACTATTTTAGTGAACAAATGTCTTGTGGCAAGGCTGTTTGGTAAAG	1070
Qy	824	tgataagggagaccccttcgtcgtacgtgcacatgta-----cgaattaaaggctttaa	877
Db	1071	TGTATTAAGGAGTCTTCCAGGTCAGACAGTATTAAGATTGCAGTGAAGACATATCA	1130
Qy	878	aagaaagaaacccaggtgtgtgagctgcagttccaacagaggttgaatattgcatg	937
Db	1131	ATGTGGAGAGCATGAAGGGGAAATGGCTTCTTATGAACAAATTGAATGATTAAGCATGG	1190
Qy	938	ctgtgacatcgaaatctcttcgtcgtacagtggtttctgcatactgacacaaagcgcctc	997
Db	1191	CTGTGCACAAACAAATTTAAGGTTGATGATGAGATCTGTACAAACCAACAGAGAGGCTCT	1250
Qy	998	ttgtatalcatacacygctaagtgaagtgttgctgtgtttaagagagcgcagccat	1057
Db	1251	TGGTCTACCCCTTTCATGGAACATCTTAGTGTGCTTCTCGTTAAGAGATTTAAATTA	1310
Qy	1058	cagaacccccccttgatctgacacacaggaagaagagttcacatagatctcttaggggc	1117
Db	1311	ATGAACCGATTTAGATTGGTCTTCAAGAAATGGGAATTTCTTTGGTGTGCCGTGGTT	1370
Qy	1118	ttctctattgcatgaccattgtgatcccaagatalcacatcgatgvtgaaagctgcaa	1177
Db	1371	TGGAATTAACCTTCATGACACATGCAATGCCAAGATCATCCACGCTACGTAAGGCTGCA	1430
Qy	1178	ataatacttggacgaagaatttggagcgctgttgatgtatttggattagctggtccca	1237
Db	1431	ATGTCTCTTGTATGGGAGACTCTCCAAACCAATGGTAGGGATTTGGGTAAAGCAAGTGA	1490
Qy	1238	tgagattaaagaagatacccatgcttacacactgctlaaggggtacacttgggtcatatagtc	1297
Db	1491	TGGACATAGGGAGGAATTAAGTAAACAAACAGGGTTGCTGGACTATGGGCCCATAGTCTC	1550
Qy	1298	ccgagtaacctgcacatcggtgaagaagtcatcagagaagacgagtcttggttatggtatca	1357
Db	1551	CTGAGTAAATTAAGACAGGAGGAGGCCATCACTGAAGACAAATATCTATGGATATGGTGTCA	1610
Qy	1358	tgaccttgagctaatctacttgagacaagagactttt-----atcttgcgtccttgcga	1411
Db	1611	TGCTTTTAGAATTTGTGACAGAGAGGAGGTGCATATGCATTCATCCTGACCTTATGAGAG	1670
Qy	1412	acgatgatgatgtatgttgttggattgggtttaaagccctttgaaagagaaaagtgtg	1471
Db	1671	AAGCGGGGAGATATATGCTCATTTGATAGGTCAAGCTATGATGATGAGAGAGGCCATATC	1730
Qy	1472	agatgtgtgtgcatcctgcacacttgtagaacaattacattgacacagaagtgttagcgctta	1531
Db	1731	TGGACTAGTAGATTCGCAATCTAGGTGGCTGTTCACAACCTTGAGAGGCTGGGAAAGGTTA	1790
Qy	1532	ttcaagtagcatctacgtctgtacccaggttgcgtcaattggagcgccaaagtgtlcaag	1591







Matches 675; Conservative 0; Mismatches 404; Indels 150; Gaps 3;

```

QY 707 cagctgaagagagaccgaagtgacaccttggtcaactgaagaaggtttctctcgcaaat 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74141 CAGGTGAAGTGGACAGAGAGATGCTTTGACAGAGTTGAGAGATTTGCATGGAGAGAC 74082

QY 767 tgcagctgcaacggatactttagt-----accatacttgaaagagtgatgtgta 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74081 TTCAGTTGGCTACAGATGAGTCAAGTAAAGAAATGTTCTCGACAGAGAGCTTTGGCA 74022

QY 821 aggtgtataaaggagacgttgcatagtctactgttagcgttaaaaggcttaaaag 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74021 AAGTTTACAAAGAGTGGCTTTGCGATGGCACCAAGTCCTGTAAGAAAGATTGACTGATT 73962

QY 881 aacgaacacccaggtggtgagctgacgttccaacagaggtggaatgattagcagctg 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73961 TTGAAGCTCCAGAGAGAGATGATAGCTTTCCAGAGAGAAATGAGATGATTAAGTTACTGTG 73902

QY 941 tgcaggaatctctctgctctagctggtttctgcataagacacagagagcgttcttg 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73901 TTCAATAGGAATCGCTGCTTATCGCTTTATCGCTTTTACACCAAGTGAACGACTTTTGG 73842

QY 1001 tatatccatacactgctaaatggaagtgtgctgctgtttag----- 1044
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73841 TGTATCTTTTCATGAGATCTAAGTGTTCATATTGCTTAAAGAGGTAAAGAAAGAGACA 73782

QY 1044 -----agag 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73781 TAAAGCTATTAACCTTTTGAATTCAGAGTTTAAAGATTTTGATCTAATACTATGCGAG 73722

QY 1048 cgtcacccatcagaacacctcccttattgccaactagaagagattgacataagctc 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73721 ATTTAAACCGGGGAGCCAGTTCGTGATGCTGATGCTGACAGAGAGAAACAGATTCGCTTAGCGCA 73662

QY 1108 gctagggggtcttcttattgcaatgacattgtatcccaagattatccatcgtatgta 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73661 GCACAGAGGCTCGAATATCTTCATGATGACCAACCGAAGATCATACAGAGAGATG 73602

QY 1168 aagagctgcaatatattttgacgaagaattgagagctgtgtgtaggtatttggtta 1227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73601 AAGAGTGCAGAAATGTTTACTAGATTAAGAACCTTTGACAGAGGTTGATGTTTGGTTTA 73542

QY 1228 gctagggctcattgaltacagaagataccatttcaactgctgaaggggtacctggc 1287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73541 GCCAAGTTGGTATGATGATGAGAGACTATGATTAACCACTCAGTGCCAGAGAACATGGGT 73482

QY 1288 tacatagctcccgagtgactctcgactggaagatcatagagaagacgagatgcttgg 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73481 CATATTGCCACGAATGATATCCACAGGAAATGCTAGGAAACCGATGTTTCCGG 73422

QY 1348 tatgggattatgctctttagagctcattcagacagagagctttagtctgtcgctt 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73421 TACGGATATATGCTTCTGAGCTTGTAACTGACAAAGAGCAATTTGATCTCGCGGTTA 73362

QY 1408 gcgaaacgagatgagtgatgtgtgtgta----- 1437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73361 GAGGAGAGAGATGATGCTTATGCTTACAGCAATGTAAGTAAATGTTTCTTTTGGCTTC 73302

QY 1437 -----ttgggttaaaagcccttga 1456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73301 TTTTCAAGCTCTCATATATGTTAAATTAACATCTTAGGTTTCAGGTGAGAAATGGA 73242

QY 1457 aagagaaaaaagtttgagatgctggtcgactctgaccttggaagaaacaattacattga 1516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73241 GAGAGAGAGATTAGACATAGTATGAGTAAGAACTTGATGAGCAATTAATAAGAGAG 73182

QY 1517 aagttgagaagcttattcaagtagcatctactgtacaccagaggttcgccaatgagcggc 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73181 AAGTTGAATATGATAGTATGAGTATGCTGCTATGACACAGAGAGACCGGAAAGAGAG 73122

QY 1577 ctaagagtgcaagagtgatcggaatgcttgaaagtgtatgacctgacagaaaagtggagc 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73121 CAGCGATGTCGAGAGTATGATGAGAAATGCTAAGAGGAGAGGCTTGACAGAGATGGAGAG 73062

```

```

QY 1637 agtgcacaaagtagaagtcataccatcaagacgtagaattagctccacatcgactctg 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73061 AGTGGCAGAAATCTGAGTGAAGTGCAGAGCAAGAGATTTGACAGAGTTTG 73002

QY 1697 aatgatacttagactcgacagataacttgcatagtctttaaattatctgctcaagataa 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73001 ATTGGGAGTGAAGATTCATTAATATCAATGATGATGATGATGATGATGATGATGATG 72942

QY 1757 cagcatataaaatgtgaatgaatata 1785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72941 AACAAAAAATCTTTTGTATCAAAAGATA 72913

```

RESULT 12  
AB025639/c  
LOCUS  
DEFINITION  
Arabidopsis thaliana genomic DNA, chromosome 3, pl clone: MML2,  
complete sequence.  
ACCESSION  
AB025639  
VERSION  
AB025639.1 GI:4589445  
KEYWORDS  
HTG.  
SOURCE  
Arabidopsis thaliana (strain:Columbia) DNA, clone:lib:Mitsui P1  
clone:MML2.  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsids.  
REFERENCE  
1 (sites)  
AUTHORS  
Nakamura,Y.  
TITLE  
Structural Analysis of Arabidopsis thaliana Chromosome 3. II  
JOURNAL  
Unpublished (1999)  
AUTHORS  
Nakamura,Y.  
TITLE  
2 (bases 1 to 84896)  
Direct Submission  
Submitted (02-APR-1999) to the DDBJ/EMBL/GenBank databases.  
Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of  
Gene Structure 2; 1532-3, Yana, Kisarazu, Chiba 292, Japan  
(E-mail:y.nakamura@kazusa.or.jp, Tel:+81-438-52-3935,  
Fax:+81-438-52-3934)

FEATURES  
Location/Qualifiers  
source  
1..84896  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="MML2"  
BASE COUNT 27260 a 15039 c 15617 g 26980 t  
ORIGIN

Query Match 13.6%; Score 246.8; DB 7; Length 84896;  
Best Local Similarity 61.6%; Pred. No. 3.1e-45;  
Matches 471; Conservative 0; Mismatches 222; Indels 71; Gaps 2;

```

QY 707 cagctgaagagagaccgaagtgacaccttggtcaactgaagaaggtttctctcgcaaat 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67675 CAGAGCAAAACAGAGAGAAATGCTGCTAGGAAATCAAGAGATTAAATTTCAAGAAC 67616

QY 767 tgcagctgcaacggatactttagtaca-----tacttgaagaagtgatgttgta 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67615 TTCAATTCGCAACTGATTAACCTTCAGCAGCAAGAAATCTGTCGAGAAAAGAGGTTTGGAA 67556

QY 821 aggtgtataaaggagacgttgcatagtctactgttagcagatgtaaaagccttaagaag 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67555 ATGCTATTAAGAGTGTGCTTCTATGATGAGATGATGATGATGATGATGATGATGATGATG 67496

QY 881 aacgaacacccaggtggtgagctgacgttccaacagaggtggaatgattagcagctg 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67495 TAAACAAATGCTGTGAGAGTTCAGTTTCAAGACAGAGCTTGAAGAAATGATTAAGCTTGGCG 67436

```



REFERENCE	1 (bases 1 to 134784)	
AUTHORS	Washington University Genome Sequencing Center.	
TITLE	The A. thaliana Genome Sequencing Project	
JOURNAL	Unpublished (1997)	
REFERENCE	2 (bases 1 to 134784)	
AUTHORS	Sun, H., Woldmann, P., Johnson, D. and Gibson, A.	
TITLE	The sequence of A. thaliana Fl17123	
JOURNAL	Unpublished (1999)	
REFERENCE	3 (bases 1 to 134784)	
AUTHORS	Waterston, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (17-JUN-1999)	Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	Submitted by:	

**MAPPING:** Clones were assigned to the YAC map by hybridization by M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, WashU, to pick the best candidates for sequencing.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems such as compressions and repeats; all regions were covered by sequence from more than one subclone

[illegible]



[illegible]



query Match

11.9%; Score 215.4; DB 8; Length 120787;

	Best Local Similarity	58.2%;	Pred. No. 3.3e-38;	
	Matches	479;	Conservative	0; Mismatches 251; Indels 93; Gaps 2;
Qy 701	atctgccagctgaagaagaccagaagtgcaccttgctaacctgaagaaggtttctctgc	760		
Db 109092	ATTTTGCACAAACAAATATGATCCGAAGTGAATTAGGGCACTTGAAGAGGTATACATTCA	109033		
Qy 761	gagaattgcgaatgcgaagatactt-----tagtacatacttggaaagagtgtgat	814		
Db 109032	AAGAAGCTTGATCTGCGACACCAATATTTCAACCTCGAAGAACATCTCCGGAAGGGGGAT	108973		
Qy 815	tggtaagtgatataaaggaagcgctctgctatgtgctactttgaagcaagttaaaggctta	874		
Db 108972	ACGGGATTTGTGTACAAAGGACACTTAAACGATGGAACTTTGGGGGTGTCAAACGCTCTCA	108913		
Qy 875	aagaagaacgaacacacaggtgtgtagctgcaggttccaacagaggtggaaatgatga	934		
Db 108912	AGGACTGTACATTTGCGGGTGGAGAGTCCAGTTTCAGACAGAAAGTAGACATTAAGTT	108853		
Qy 935	tggctgtgcacgaactctctgcgtctacgttggttcttcgatgacacacagagcgagc	994		
Db 108852	TGGCTTTTATCCCAATCTCTCCGGCTCCCGGCTTCTGTGAGACGAACACAGAGAGAA	108793		
Qy 995	tctctgtatcatcatcaacaggtctaaggaaagtgtgtgtgttttaaggagacgctcagc	1054		
Db 108792	TTTTAGCTTACCTTATCATGCCAAATGGAGATGTGCATCAGCTTTAAAGATTAATGCA	108733		
Qy 1055	catc-----	1059		
Db 108732	TTTTCCAAAATAACTGCTAATTTCTTCACAGCTTTTTCAGTAAGTTTAAATAGTTCATCTT	108673		
Qy 1059	-----agaactccctctgattggtccactagga	1087		
Db 108672	GCTACATTTTTTGTGTAGATAATATCCGTGGAGACCAAGCTTGAAGTGTGCAAGAAAG	108613		
Qy 1088	agaagattgcacagagatctgctaggggctttcttatttgcagacattgcatccca	1147		
Db 108612	AGAAGATACGGGTGTGGGACAGCGAAGGACATAGTTTACTACAGACATGTGACCCGA	108553		
Qy 1148	agattatcatcgttgatgtlaaaagctgcacaatataatttggacgaagaatttggagctg	1207		
Db 108552	AGATTATACCCCGATGTGAAGACAGCTAACATCTGTTGATGAGAGACTTGTGAAGCAG	108493		
Qy 1208	tttgaagtgttttgggttagctgaagctcatgattataaagaataaccatggttacaagt	1267		
Db 108492	TTGTGGTGAATTTGGGTATGCTTAACCTTTAGACCAATAGAGACTTCATGTCACAACTG	108433		
Qy 1268	cttgaaggggttaccttgggtctacatagctcccgagtaacctctgcagtgygaagtcacag	1327		
Db 108432	CAGTCCGTGGACATGTTGGGCAATATGCACCTGATTAATTCACAGGGGTACAGCTCAAC	108373		
Qy 1328	agaagaacgagatcctttgtgtataggaatatgctctttagaactattacagtggaagag	1387		
Db 108372	AGAAGACTGATGATCTTTGGCTTTGGCATATCTCTCTTGACTCAATTAATGTGTAGAAAG	108313		
Qy 1388	ctttgatctgtcgcgcttgcgaacgatgatgatgtatgtt 1430			
Db 108312	CTCTTGATTTTGGCAATCCGCACACACAAAGAGGTAAATGCT 108270			
RESULT 15	AC005170	110211 bp	DNA	PLN 17-DEC-1999
LOCUS	AC005170/c			
DEFINITION	Arabidopsis thaliana chromosome II section 135 of 255 of the			
ACCESSION	AC005170	AE002093		
VERSION	AC005170.2	GI:6598448		
KEYWORDS	HTG.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	euphyllophytes; Spermatophyta; Magnoliophyta; endicotsyledons; core			
	endicots; Rosidae; eurosids II; Brassicales; Brassicaceae;			



REFERENCE 1 (bases 1 to 110211)  
 AUTHORS Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldlyum, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Rongling, C.M., Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., VanAken, S.E., Unayam, L., Tallon, L.U., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhagen, G.P., Preuss, D., Nieman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C.  
 TITLE Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*  
 JOURNAL Nature 402 (6763), 761-768 (1999)  
 MEDLINE 20083487  
 REFERENCE 2 (bases 1 to 110211)  
 AUTHORS Lin, X.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
 COMMENT On Dec 17, 1999 this sequence version replaced gi:3738313. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (<http://www.tigr.org/tldb/at/at.html>).

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://artur.epm.ornl.gov/pub/xgrail>), GeneFinder (Phil Green, University of Washington), GENSCAN (Chris Burge, <http://gnomic.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curation based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted as tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to the bottom of the chromosome.

We thank the GSH/Maxhu/ABI consortium for sequencing BAC clones FEP23, F506, T17A5, and T13L16, the BSA group for sequencing clone F13D3, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Peng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: [atetgr.org](mailto:atetgr.org).

FEATURES  
 source Location/Qualifiers  
 1..110211  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="2"  
 misc\_feature complement(1..2793)  
 /note="Sequence from clone F27L4"  
 repeat\_region complement(356..401)  
 /rpt\_family="(TAA)n"  
 /rpt\_family="(TAA)n"  
 complement(join(1349..2058,2731..3012))  
 /gene="At2g23810"  
 complement(1549..3012)  
 /gene="At2g23810"

CDS  
 /note="F27L4.1"  
 complement(join(1744..2058,2731..3003))  
 /gene="At2g23810"  
 /codon\_start=1  
 /product="similar to senescence-associated protein"  
 /protein\_id="AAPI8611.1"  
 /db\_xref="GI:5598449"  
 /translation="MFLILLVFCITVFPAFVTKNGAGEHIEGKYEKKLDYSTWLQKREVNKRNKINISCLVESKVCISLEAKFVNPNVSVFYKEHITLALSGCCRPDSCEGFEYVNPVTWETKNTGTETNPDCOTMNAKKEKLFDCQSCAGILLDNVKSAMKVAIVNIVPLVFLIVYVSCAFRRNNKRDSDYSRTYGYKP"  
 2794..110211  
 /note="Sequence from clone T29E15"  
 complement(3567..3677)  
 /rpt\_family="(TAA)n"  
 /rpt\_family="(TAA)n"  
 complement(3570..3641)  
 /rpt\_family="(TAA)n"  
 complement(4565..4612)  
 /rpt\_family="(CA)n"  
 6379..6484  
 /rpt\_family="(CAAAA)n"  
 join(<6480..6749,7602..7726,7811..7848,7929..8039,8152..8268,8487..8786)  
 /gene="At2g23820"  
 <6480..8786  
 /gene="At2g23820"  
 /note="T29E15.2"  
 join(6480..6749,7602..7726,7811..7848,7929..8039,8152..8268,8487..8557)  
 /gene="At2g23820"  
 /note="unknown protein"  
 /codon\_start=1  
 /protein\_id="AAC63656.1"  
 /db\_xref="GI:3738315"  
 /translation="MKVNRVRFKKSILVFLVPSFTKRIKRTTCAAGASSPNRAIHCAASDPQSGSVSSPPNVAAVSSSSASSASADLSLCTRLTTPAGWIKVDYKDPEIADHMYRMGIMLISDDIPGVSRDCKMAIVHDIAEIVDDIPSCISKEEKNRRSESLAEHMKLIGGERAKETIALMREYENSSPEAKVVKDFKVELLIDLALEYEGCKFOTNIGKAMASSEIYSRRRKH"  
 complement(6654..6712)  
 /rpt\_family="(GAA)n"  
 join(<9531..9582,9709..9779,9866..9947,10054..10169,10268..10396)  
 /gene="At2g23830"  
 <9531..10396  
 /gene="At2g23830"  
 /note="T29E15.3"  
 join(9531..9582,9709..9779,9866..9947,10054..10169,10268..10396)  
 /gene="At2g23830"  
 /note="unknown protein"  
 /codon\_start=1  
 /protein\_id="AAC63657.1"  
 /db\_xref="GI:3738316"  
 /translation="MSNNLEIEIPMYLQPFPEELKKQMSGLITNTKDNVAFKYTKRNKYCVAPNGLILPSTCKVLTWQAKQEVSDMQSEKFNISVLASPSVTAKEVIREKFSKESGVHEETKLRVYVSTTNTITSSPRTRRQFIQCFC"  
 complement(11332..11371)  
 /rpt\_family="(TAA)n"  
 15801..15917  
 /rpt\_family="(TAA)n"  
 15853..15918  
 /rpt\_family="(TAA)n"  
 /rpt\_family="(CAAAA)n"  
 16051..16136  
 /rpt\_family="(TAA)n"  
 complement(join(<20442..20537,20617..20691,20894..20948,21145..221485))  
 /gene="At2g23840"  
 complement(<20442..221485)  
 /gene="At2g23840"  
 /note="T29E15.4"  
 complement(join(20442..20537,20617..20691,20894..20948,21145..221485))







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:48:42 ; Search time 446.21 Seconds  
(without alignments)  
1017.679 Million cell updates/sec

Title: US-09-180-798-2

Perfect score: 1815  
Sequence: 1 gacaataaccattgaatat.....ttcaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : N.Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1803	99.3	1814	1	Daucus carota SERK
2	933.2	51.4	2089	1	Arabidopsis thaliana
3	560.2	30.9	6695	1	Daucus carota SERK
4	270.8	14.9	4081	1	Arabidopsis thaliana
5	143.4	7.9	3842	1	Tomato Xa21 clone
6	142.4	7.8	4104	1	Arabidopsis thaliana
7	142	7.8	3045	1	Maize Xa21 gene DT
8	140	7.7	1554	1	Tomato RRK gene cl
9	140	7.7	3293	1	Arabidopsis thaliana
10	138.4	7.6	3176	1	Arabidopsis thaliana
11	121.2	6.7	2640	1	S-receptor
12	115.4	6.4	2833	1	S-receptor kinase
13	114.8	6.3	2685	1	SRK-A10 CDNA. DNA
14	114.8	6.3	2685	1	SRK-A10 gene mutan
15	112.6	6.2	9295	1	Arabidopsis thaliana
16	110.4	6.1	2749	1	Brassica sp. S-rec
17	107.2	5.9	2780	1	Arabidopsis thaliana
18	103.4	5.7	788	1	Arabidopsis thaliana
19	103.4	5.7	1063	1	Arabidopsis thaliana
20	101.8	5.6	981	1	Arabidopsis thaliana
21	99.6	5.5	2443	1	Tomato Pto protein
22	99.4	5.5	1269	1	ZmPK1 homologue pr
23	95.4	5.3	894	1	Arabidopsis thaliana
24	92	5.1	2704	1	Arabidopsis thaliana
25	88.8	4.9	2749	1	S-receptor kinase
26	87.6	4.8	5733	1	Arabidopsis thaliana
27	84.2	4.6	1106	1	Arabidopsis thaliana
28	76	4.2	960	1	Arabidopsis thaliana
29	73	4.0	3573	1	Protein kinase gen
30	73	4.0	6471	1	Tomato pathogen re
31	70.2	3.9	3979	1	CF-5 pathogen resi
32	70.2	3.9	3979	1	CF-5 pathogen resi
33	70.2	3.9	4123	1	CF-5 pathogen resi
34	64.6	3.6	3541	1	CF-5 pathogen resi

35	62.6	3.4	19639	1	X23524	O. longistaminata
36	59.6	3.3	7204	1	X23527	O. longistaminata x
37	58.6	3.2	277	1	V60314	Resistance marker
38	58.6	3.2	3921	1	T31300	Rice Xa21 disease
39	58	3.2	13340	1	X23522	O. longistaminata
40	57.2	3.2	8416	1	X23523	O. longistaminata
41	55.6	3.1	5940	1	X23526	O. longistaminata x
42	54.8	3.0	2075	1	T49435	Tomato polygalactu
43	54.2	3.0	960	1	T35826	Protein kinase gen
44	54	3.0	9424	1	X23525	O. sativa Xa21 gen
45	53.8	3.0	6256	1	T31299	Rice Xa21 disease

## ALIGNMENTS

RESULT	ID	Score	Length	Description
1	V06571	99.3%	1814	Daucus carota SERK
2	V06571	99.3%	1814	Daucus carota SERK
3	V06571	99.3%	1814	Daucus carota SERK
4	V06571	99.3%	1814	Daucus carota SERK
5	V06571	99.3%	1814	Daucus carota SERK
6	V06571	99.3%	1814	Daucus carota SERK
7	V06571	99.3%	1814	Daucus carota SERK
8	V06571	99.3%	1814	Daucus carota SERK
9	V06571	99.3%	1814	Daucus carota SERK
10	V06571	99.3%	1814	Daucus carota SERK
11	V06571	99.3%	1814	Daucus carota SERK
12	V06571	99.3%	1814	Daucus carota SERK
13	V06571	99.3%	1814	Daucus carota SERK
14	V06571	99.3%	1814	Daucus carota SERK
15	V06571	99.3%	1814	Daucus carota SERK
16	V06571	99.3%	1814	Daucus carota SERK
17	V06571	99.3%	1814	Daucus carota SERK
18	V06571	99.3%	1814	Daucus carota SERK
19	V06571	99.3%	1814	Daucus carota SERK
20	V06571	99.3%	1814	Daucus carota SERK
21	V06571	99.3%	1814	Daucus carota SERK
22	V06571	99.3%	1814	Daucus carota SERK
23	V06571	99.3%	1814	Daucus carota SERK
24	V06571	99.3%	1814	Daucus carota SERK
25	V06571	99.3%	1814	Daucus carota SERK
26	V06571	99.3%	1814	Daucus carota SERK
27	V06571	99.3%	1814	Daucus carota SERK
28	V06571	99.3%	1814	Daucus carota SERK
29	V06571	99.3%	1814	Daucus carota SERK
30	V06571	99.3%	1814	Daucus carota SERK
31	V06571	99.3%	1814	Daucus carota SERK
32	V06571	99.3%	1814	Daucus carota SERK
33	V06571	99.3%	1814	Daucus carota SERK
34	V06571	99.3%	1814	Daucus carota SERK
35	V06571	99.3%	1814	Daucus carota SERK
36	V06571	99.3%	1814	Daucus carota SERK
37	V06571	99.3%	1814	Daucus carota SERK
38	V06571	99.3%	1814	Daucus carota SERK
39	V06571	99.3%	1814	Daucus carota SERK
40	V06571	99.3%	1814	Daucus carota SERK
41	V06571	99.3%	1814	Daucus carota SERK
42	V06571	99.3%	1814	Daucus carota SERK
43	V06571	99.3%	1814	Daucus carota SERK
44	V06571	99.3%	1814	Daucus carota SERK
45	V06571	99.3%	1814	Daucus carota SERK



QY 121 aattacatgcaatcactgagcttacccttgacaataatgggttcttaagaattggag 180  
DB 121 aattacatgcaatcactgagcttacccttgacaataatgggttcttaagaattggag 180  
QY 181 cttaacagaataaataagatgagcaattccttagtgaattctgggaattcacaatttg 240  
DB 181 cttaacagaataaataagatgagcaattccttagtgaattctgggaattcacaatttg 240  
QY 241 gtagagcttggacccttaataatgagcttctcgtgaccctataccggacataaggaaag 300  
DB 241 gtagagcttggacccttaataatgagcttctcgtgaccctataccggacataaggaaag 300  
QY 301 cttaacagaataaataagatgagcttctcgtgaccctataccggacataaggaaag 360  
DB 301 cttaacagaataaataagatgagcttctcgtgaccctataccggacataaggaaag 360  
QY 361 tcaactgacataatcaacaactcctcaagtcctggaattcacaacaactcggactaaggaa 420  
DB 361 tcaactgacataatcaacaactcctcaagtcctggaattcacaacaactcggactaaggaa 420  
QY 421 ccaagtaacgagataatgagcttctcgtgaccctataccggacataaggaaag 480  
DB 421 ccaagtaacgagataatgagcttctcgtgaccctataccggacataaggaaag 480  
QY 481 aattatgtgagcccttaacatgagagccctcgtgaccctataccggacataaggaaag 540  
DB 481 aattatgtgagcccttaacatgagagccctcgtgaccctataccggacataaggaaag 540  
QY 541 cctcgttcaatcccaatcaacagtaacagctcgaagcaaaatggtcccaatggagct 600  
DB 541 cctcgttcaatcccaatcaacagtaacagctcgaagcaaaatggtcccaatggagct 600  
QY 601 attgagggggagtgagcttctcgtgaccctataccggacataaggaaag 660  
DB 601 attgagggggagtgagcttctcgtgaccctataccggacataaggaaag 660  
QY 661 gcatgtgagggaggaagaaacccggagaaacattcttgatgtgacagtgaaagagac 720  
DB 661 gcatgtgagggaggaagaaacccggagaaacattcttgatgtgacagtgaaagagac 720  
QY 721 ccaagaaatgagcttctcgtgaccctataccggacataaggaaag 780  
DB 721 ccaagaaatgagcttctcgtgaccctataccggacataaggaaag 780  
QY 781 gatacttgaatacactgagaaaggttgagatttgtaagtgatcaagaagcagctt 840  
DB 781 gatacttgaatacactgagaaaggttgagatttgtaagtgatcaagaagcagctt 840  
QY 841 gctgagtgagcttctcgtgaccctataccggacataaggaaag 900  
DB 841 gctgagtgagcttctcgtgaccctataccggacataaggaaag 900  
QY 901 ctgagcttcaacagaagtgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 960  
DB 901 ctgagcttcaacagaagtgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 960  
QY 961 ctgagcttcaacagaagtgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 1020  
DB 961 ctgagcttcaacagaagtgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 1020  
QY 1021 gtagagtgagcttctcgtgaccctataccggacataaggaaag 1080  
DB 1021 gtagagtgagcttctcgtgaccctataccggacataaggaaag 1080  
QY 1081 actagagaagagatgcacactgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 1140  
DB 1081 actagagaagagatgcacactgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 1140  
QY 1141 gatcccaagagatgcacactgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 1200  
DB 1141 gatcccaagagatgcacactgagaaatgagtgatgagcttctcgtgaccctataccggacataaggaaag 1200  
QY 1201 gagagcttctcgtgaccctataccggacataaggaaag 1260

DB 1201 gagagcttctcgtgaccctataccggacataaggaaag 1260  
QY 1261 acaactgctgaagggagcttctcgtgaccctataccggacataaggaaag 1320  
DB 1261 acaactgctgaagggagcttctcgtgaccctataccggacataaggaaag 1320  
QY 1321 tcatcagaagagagagcttctcgtgaccctataccggacataaggaaag 1380  
DB 1321 tcatcagaagagagagcttctcgtgaccctataccggacataaggaaag 1380  
QY 1381 cagagagcttctcgtgaccctataccggacataaggaaag 1440  
DB 1381 cagagagcttctcgtgaccctataccggacataaggaaag 1440  
QY 1441 gttaaagccttctcgtgaccctataccggacataaggaaag 1500  
DB 1441 gttaaagccttctcgtgaccctataccggacataaggaaag 1500  
QY 1501 aattacattgacagagagtgagagcttctcgtgaccctataccggacataaggaaag 1560  
DB 1501 aattacattgacagagagtgagagcttctcgtgaccctataccggacataaggaaag 1560  
QY 1561 tgcgaatgagagagagtgagagcttctcgtgaccctataccggacataaggaaag 1620  
DB 1561 tgcgaatgagagagagtgagagcttctcgtgaccctataccggacataaggaaag 1620  
QY 1621 gtagaagagtgagagagtgagagcttctcgtgaccctataccggacataaggaaag 1680  
DB 1621 gtagaagagtgagagagtgagagcttctcgtgaccctataccggacataaggaaag 1680  
QY 1681 ccaactgagagagtgagagcttctcgtgaccctataccggacataaggaaag 1740  
DB 1681 ccaactgagagagtgagagcttctcgtgaccctataccggacataaggaaag 1740  
QY 1741 tctggtcgaagagtgagagcttctcgtgaccctataccggacataaggaaag 1800  
DB 1741 tctggtcgaagagtgagagcttctcgtgaccctataccggacataaggaaag 1800  
QY 1801 aaaaaaa 1815  
DB 1801 aaaaaaa 1815  
RESULT 2  
V06591  
ID V06591 standard; cDNA to mRNA; 2089 BP.  
AC 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
OS Arabidopsis thaliana.  
FH key location/Qualifiers  
FT CDS 195..2072  
FT /tag= a  
FT /product= SERK protein  
PD W09743427-A1.  
PD 20-NOV-1997.  
PE 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOV) NOVARTIS AG.  
PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;  
DR WPI; 98-086529/08.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 27; Pages 91-95; 123pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds.  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing



CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.

Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;

Query Match 51.4%; Score 933.2; DB 1; Length 2089;  
Best Local Similarity 74.4%; Pred. No. 1.6e-229;  
Matches 1191; Conservative 0; Mismatches 403; Indels 6; Gaps 1;

QY 175 ttgagcttacagcaatacaataagtgagcaattccctagatcttggaatgaca 234  
DB 486 TTGGAGCTTTACAGTAACAACATTAAGTGGCCGATTCCTAGTAATCTTGGAAATCTGACA 545  
QY 235 aatttgagcttgagcctatacatgaatagcttctctggaactataccgagacatia 294  
DB 546 AACTAGTGTGTTGGATCTTTACTTAACAAGCTTCGCGCTCTTCGGAATCATTTG 605  
QY 295 ggaagcttcaagagcttaagatcttgcgtctcaacaagaacgctctctgccaatt 354  
DB 606 GGAAGGCTTTCACAAAGCGAATTCCTCCGCTTACACAAACAGTCTCACTGGGTCAAT 665  
QY 355 ccaatgactgactgaataatacaactctcaagctcctgagattatacaacaatgaca 414  
DB 666 CCRATGCACTGACCAATATTAATACCTCAAGTGTAGATATTAACAATAACACATC 725  
QY 415 tcaagagacagtaacggaataagctcaattcttctgtttacacatacagtttgcaat 474  
DB 726 TCTGTGATCACTCTCGTCAATAGGCTCCTCTCACTCTACACCCATCAGTTTGGTAAT 785  
QY 475 aattgaattatgtgagaccgtaactgaagcgcctgagccttgatctccccaatttg 534  
DB 786 AACTTACACTATGTGACCTGTACAAAGTCAACCCATGCTCGGATCTCCCGCTTTCT 845  
QY 535 ccaccacccgttcaatcccaatacaacagtaacagctccaggaacaaatgtgccat 594  
DB 846 CCTCCACACACTTTTATTAACAACCTCCGCCAGTTTCCACCCGAGTGGATATTAAC 905  
QY 595 ggaactatgctggggagtagctgctgctgcttactgcttgcagactgcaatg 654  
DB 906 GGAGCAATAGCTGGTGGAGTGGCTGACAGTGGCTTGGCCCTTGGTGGCTCTCGCAAT 965  
QY 655 gcaattgcatggtggcgagagaagaacccggagagaacattcttgaatgtgcaagctgaa 714  
DB 966 GCCCTTCTGCTGGTGGCGACCAAGAGGCCCACTAGATATTTCTTCGATGCTCCCTGCCGA 1025  
QY 715 gagagaccagaagtgcaccccttggtcaactgaagaggtttctctgcgaagatgcaagtc 774  
DB 1026 GAAGATTCACAAGATCTATCTGGGACACCTCAAGAGGTTTCTTGGCGGAGCTACAAAG 1085  
QY 775 gcaacgatactttagt-----accatacttggaagaggttgattgtgtaagtgat 828  
DB 1086 GCGAGTATGGGTTTACTTAACAAGAACATTTTGGGCGAGAGTGGGTTGGGAAAGTCTAC 1145  
QY 829 aaggagagccttgctgtagctcaactgtgtagcagtttaaaaggtttaaagaagaaca 888  
DB 1146 AAGGAGAGCTTGGGACGGAACCTTGTGGCTGTCTCAAGAGACTGGAAGGAAAGACGAACT 1205  
QY 889 ccaaggtgtgagctgcagtttcaacaagaggttgaaatgattagcagtgctgcacatga 948  
DB 1206 CCAAGTGGAGAGCTCCAGTTTCAACAAGAGATGATTAAGTATGGAGTTTCATCGA 1265  
QY 949 aatcttctgctacagtggttctgcatgacacacagaagcggtcttctgtatatacca 1008

DB 1266 AACCTGTTGAAATACAGAGGTTTCTGATGACACCGCACAGAGATGCTTGTGATCC 1325  
QY 1009 taactggttaagtaagtggtgctgtgtttaagagaaggtcaacataagaacccccc 1068  
DB 1326 TACATGGCAATGGAAGTGTGCTTCTGTCTCAGAGAGGCGCACCGTACAAACCTCCG 1385  
QY 1069 ctgattggccaactagaagaagattgcaactagatcctctgaaggggtcttctattg 1128  
DB 1386 CTGTATTGGCCAAAGCGGAAAGAGATGGCGCTAGCGCTCAGCTGAGGTTGTCTACCTA 1445  
QY 1129 catgacacatgtatcccaagattatccatcogtgaatgtaaaagctgcaatatattg 1188  
DB 1446 CATGATCACTCGATCCGAAGATCATTCACCGTACGCTTAAGAGCAACATCTCTTGA 1505  
QY 1189 gacgaagattgagctgtgtgtagtatttggtgattagctgagctcagatgataag 1248  
DB 1506 GACGAAGATTCGAGGCGGTGTGGAGATTTCCGGTTGGCAAGCTTATGACTATAAA 1565  
QY 1249 gataccatgttaacaactgctgtaagggttaaccttggtctacatagctcccgatctc 1308  
DB 1566 GACACTCAGTACAGACAGAGCTCCGTGGCACATCGGTCACTCGCTCCAGATATCTC 1625  
QY 1309 tgcactggaagtcatacagaagaacgagatgctcttggttaaggatlatgctttagag 1368  
DB 1626 TCAACCGGAAATCTTCAGAGAAACCGAGCTTTCCGATACGAATCATGCTCTAGAA 1685  
QY 1369 ctcaattactgacagagagcttctgactctgctgctcttgagacagatgattgattg 1428  
DB 1686 CTAACTCAGAGACAAAGAGCTTTGATCTCGCTGGGTAGCTTAACGACGAGAGCTCATG 1745  
QY 1429 ttgttgatgtgttaaaagccttctgaagaagaagaagttgagatgctgctgctcct 1488  
DB 1746 TTACTTGACTGGGTGAAGATGTGTGAAGAGAAAGATAGAGATGTAGTGAATCCA 1805  
QY 1489 gaactggaacaaattacattgacacagaagttgagagcttattcaagtagaattatc 1548  
DB 1806 GATCTTCAAAACAACTACGAGAGAGAGAACTGGAACAGATGATCAAGTGGCTTCTA 1865  
QY 1549 tgtaccagaggttgccaatgagcagcgctgaagtgtgcaaggtgagccgaatgctgaa 1608  
DB 1866 TGCACGCAAGGATCACCATGGAAGAACCAAGATGTCTGAAGTTTGAAGATCTCGGAA 1925  
QY 1609 gctgagccttgacagaagaatgagcaggtgagcaaaaagtagaagcatccatcaagac 1668  
DB 1926 GGAGATGGGCTTGGCGGAATATGGGACCAATGCCAAAAGTGTGATTTGAGGGAAGAG 1985  
QY 1669 gtgaattagctccacatcgaaactctgtaatgtaacctagatcagacataacttgat 1728  
DB 1986 ATTGATTGAGTCTAATCTTAACCTGATGATGATCTGTGATCTACTTCAATTTGCAC 2045  
QY 1729 gctttgaattatctggtccaaagataacagcatataaa 1768  
DB 2046 GCCGTTGAGTTATCTGTCCAAAGTAAAAAATAAAAAA 2085

RESULT 3  
V06570  
ID V06570 standard; DNA; 6695 BP.  
AC V06570;  
DR 03-NOV-1998 (first entry)  
DE Daucus carota SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KW plant breeding; ds.  
OS Daucus carota.  
FH Key  
FH Location/Qualifiers  
FT 3696..6620  
FT CDS  
FT /\*tag= a  
FT /note= "contains introns"  
FT 3731..3802  
FT /\*tag= b  
FT /number= 1  
FT Intron  
FT Intron 3851..3979



```
FT      /tag- c
FT      /number- 2
FT      4124. .4211
FT      /tag- d
FT      /number- 3
FT      4284. .4357
FT      /tag- e
FT      /number- 4
FT      4430. .4528
FT      /tag- f
FT      /number- 5
FT      4642. .4757
FT      /tag- g
FT      /number- 6
FT      4890. .4967
FT      /tag- h
FT      /number- 7
FT      5295. .5803
FT      /tag- 1
FT      /number- 8
FT      6197. .6339
FT      /tag- j
FT      /number- 9
FT      MO9743427-A1.
PN      PD
PN      20-NOV-1997.
PR      14-MAY-1997: E02443.
PR      14-MAY-1996: GB-010044.
PI      (NOVS ) NOVARTIS AG.
PI      De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;
DR      WP1: 98-086529/08.
DR      P-ESDB: M47013.
PT      Production of apomictic seeds - useful in plant breeding
PS      Claim 21: Pages 40-46: 123pp: English
CC      The sequence is that encoding SPRK, a putative receptor kinase.
CC      It may be used as part of a method of producing apomictic seeds
CC      comprising: (a) transforming plant material with a nucleotide
CC      sequence encoding a protein which in active form in a cell or
CC      cell membrane renders the cell embryogenic; (b) regenerating
CC      the transformed material into plants or carpel-containing
CC      plant parts; and (c) expressing the sequence in the vicinity
CC      of the embryo sac. The apomictic seeds and embryos thus produced
CC      can be developed into plant progeny. This is useful in plant
CC      breeding programs. Controllable and reproducible apomixis provides
CC      many advantages in plant improvement and cultivar development in
CC      the case that sexual plants are available as crosses with the
CC      apomictic plant. Apomixis provides for true-breeding, seed
CC      propagated hybrids and could shorten and simplify the breeding
CC      process so that selfing and progeny testing to produce and/or
CC      stabilise a desirable gene combination could be eliminated.
CC      Apomixis allows plant breeders to develop cultivars with
CC      specific stable traits for such characteristics as height,
CC      seed and forage quality and maturity.
SQ      Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match      30.9%; Score 560.2; DB 1; Length 6695;
Best Local Similarity 81.9%; Pred. No. 8e-134;
Matches 735; Conservative 0; Mismatches 18; Indels 144; Gaps 2;
```

```
Db      5980 GTTAGCTAGCCATCATGATTACAGGATACCACATGCTTACAGACTGCTGTAAGGGATACAT 6039
Qy      1284 gggctacatagctcccgagctaccctcgactggaagtcacagaagaaccgatgctc 1343
Db      6040 TGGGCACATFAGCTCCGAGTACCTTCGCACTGGAAGCATCAGAGACCGATGCTTT 6099
Qy      1344 tggcttaggatattgctcttagagctactctgtagaagaagagctttgactctgctg 1403
Db      6100 TGGTTATGGGATATGCTCTCTAGAGCTCATTTACTGACAGAGGGCTTTGATCTTGCTCG 6159
Qy      1404 ccttcgacagatgatgatgtatgtgttgatg----- 1440
Db      6160 CATTGCGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6219
Qy      1440 ----- 1440
Db      6220 GGTATATATTTACATATATATAGTCTTACTACTCTGTTGTGSCCTGTTTATTTTCC 6279
Qy      1440 ----- 1440
Db      6280 TGCCTGATTTGATTTCTTAGTCATGTTATGATATGACCTGCTTGCATGCTTTTAG 6339
Qy      1441 gttaaaagccctttgaagaagaagaagttgsgatgctggtcgcactcgactgagagac 1500
Db      6340 GTTAAAGCCTTTTGAAGAGAAAAGTTGAGATGCTGCTGATCCTGACCTGCAGAAC 6399
Qy      1501 aattacattacacagaagttgagcagcttatcagaatgacttactctgtaccagagt 1560
Db      6400 AATTACATTACACAGAGAGTTGAGCCTTATTCAATGACATTTCTCTGTACCCAGGCT 6459
Qy      1561 tcgccaatggagcgcgctaagatgctcagagtagtccgaatgcttgaagtgatggcct 1620
Db      6460 TCGCCATGAGAGCGCCTAAGATGTCAGAGTAGTCCGAATCTTGAAGGTATGCGCTT 6519
Qy      1621 gcagaagaagtcgagagtgctgcaaaaagtagaagtcattccatcagaagctagatgact 1680
Db      6520 GCAGAAAAGTGGAGCGGTGGCAAAAAGTTGAAGTCAATCAGACGTACAGATTAATTA 6579
Qy      1681 ccacatgaactcttgatgacccctagactcgacagataacttgatgctttgaatla 1740
Db      6580 CCAATGGAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6639
Qy      1741 tctgtccaaataaacaagcatataaattgtgaatgaatataatattttatgctt 1797
Db      6640 TCTGTCCAAAGATTAACAGCATATAAATGT-AATGAATTAATATTTTATGCTT 6695

RESULT 4
ID      V06585
AC      V06585;
DT      03-AUG-1998 (first entry)
DE      Arabidopsis thaliana SERK gene.
KW      receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS      Arabidopsis thaliana.
FH      Key
FT      CDS
FT      /tag- a
FT      /note= "contains introns"
FT      3731. .3802
FT      /tag- b
FT      /number- 1
FT      3851. .3979
FT      /tag- c
FT      /number- 2
FT      4124. .4211
FT      /tag- d
FT      /number- 3
FT      4284. .4357
FT      /tag- e
FT      /number- 4
FT      4430. .4528
FT      Intron
FT      Intron
```



```

FT      /tag= f
FT      /number= 5
FT      /tag= 4642. .4757
FT      /tag= 9
FT      /number= 6
FT      /tag= 4890. .4967
FT      /tag= h
FT      /number= 7
FT      /tag= 5295. .5803
FT      /tag= 1
FT      /number= 8
FT      /tag= 6197. .6339
FT      /tag= j
FT      /number= 9

PN      WO9743427-A1.
PD      20-NOV-1997.
PF      13-MAY-1997; E02443.
PR      14-MAY-1996; GB-010044.
PA      (NOVS ) NOVARIIS AG.
PI      De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
PI      WPI: 98-086529/08.
DR      P-PSDB; W47017.
PT      Production of apomictic seeds - useful in plant breeding
PS      Claim 26; Pages 64-67; 123pp; English.
CC      The sequence is that encoding SERK, a putative receptor kinase.
CC      It may be used as part of a method of producing apomictic seeds
CC      comprising: (a) transforming plant material with a nucleotide
CC      sequence encoding a protein which in active form in a cell or
CC      cell membrane renders the cell embryogenic; (b) regenerating
CC      the transformed material into plants or carpel-containing
CC      plant parts; and (c) expressing the sequence in the vicinity
CC      of the embryo sac. The apomictic seeds and embryos thus produced
CC      can be developed into plant progeny. This is useful in plant
CC      breeding programs. Controllable and reproducible apomixis provides
CC      many advantages in plant improvement and cultivar development in
CC      the case that sexual plants are available as crosses with the
CC      apomictic plant. Apomixis provides for true-breeding, seed
CC      propagated hybrids and could shorten and simplify the breeding
CC      process so that selfing and progeny testing to produce and/or
CC      stabilise a desirable gene combination could be eliminated.
CC      Apomixis allows plant breeders to develop cultivars with
CC      specific stable traits for such characteristics as height,
CC      seed and forage quality and maturity.
SQ      Sequence 4081 BP; 1120 A; 770 C; 785 G; 1406 T;

Query Match      14.9%; Score 270.8; DB 1; Length 4081;
Best Local Similarity 64.6%; Pred. No. 7.9e-60;
Matches 504; Conservative 0; Mismatches 157; Indels 119; Gaps 2;

QY      707 cagctgaagaagagaccagaagatgacactgtgcaactggaagaggtttctctgcgagaat 766
DB      3302 CACCAGAGAAAGATCCAGAGATTCATCTGGACAGCTCAAGAGGTTTCTTCTGGGAGC 3361
QY      767 tgcgaatgcgaaggaacttttagt-----acatacttgcgaagagtgtagttgta 820
DB      3362 TAAAGATGGGAGTATGGGTTAGTAAACAAGAACATTTTGGGCGAGGTGGTTTGGGA 3421
QY      821 aggtgtataagagagccctgtgctgactgtgactgtgactgtgactgtgactgtgactgtg 880
DB      3422 AAGTCTTAACAAGAGGAGCGCTTGGCAGAGAACTCTTCTCTCAAGAGACTGAAGAGAG 3481
QY      881 aacgaacaacaggtgtgctgctgcaagttccaacagaggtggaatgtagcatgtgctg 940
DB      3482 AGCGAAGCTCCAGCTGAGAGCTCCAGATTCCAACAAACAAGATGAAGATGATTAAGTATG 3541
QY      941 tgcatacgaatctctctgcgtctacgtgctgttctgcatgacacacacagagcggtcttg 1000
DB      3542 TTTCATCGAAACCTGTGATGATTACGAGGTTTCTGTTATGACACCGACGACGAGATTGCTTG 3601
QY      1001 tatatcatatcatgctgaatgtaagatgttgcgtcgtgtttaag-----1044
DB      3602 TGATCTCTACATGCGCAATGAGAACTGTGCTCTGCTCAGAGGTAAACAACTAAACA 3661

```

```

QY      1044 -----1044
DB      3662 TTAAACATCTGTGCTCTCTCAATTACTTGAAGTGAAGTGTTCATGTTTCC 3721
QY      1044 -----agagcgtacgcatcagacctcc 1067
DB      3722 TTATGGGTTCAATATTGTGGTTACACTAATGACACAGAGAGGCCACCGTCACAACTCC 3781
QY      1068 ccttgattggccaactaggaagagatgtcacatagatctgtcagggggtcttctattt 1127
DB      3782 GCTTGATTGGCCACGCGGAAGAGAAATGCGCTAGCTCAGCTGAGGTGTGCTTACCT 3841
QY      1128 gcatgacattgtgctcccaagattatccatcgatgataagatcgaatattattt 1187
DB      3842 ACATGATCTACGTGACATCCGAATATCATTCACCGAGCTAAACACAAACATCCCTT 3901
QY      1188 ggaagaaagattgagagctgtgtgtagtatttgggttagtagctagctcattgataa 1247
DB      3902 AGACGAGAAATTCGAAGGGGTTGTTGAGATTTGGGGTTGGCAAAAGCTAATGGACTATTA 3961
QY      1248 ggatacccatgttacaactgtctgtaagaggtacattggtcattacatcccgatcct 1307
DB      3962 AGACACTCAGTGACAAACGACAGTCCGTGGCACCATCGGTCCATCCCTCCAGAAATCT 4021
QY      1308 ctgcagctggaagatcatcagaagaagccgatgctgtcttggtatggattagttgcttaga 1367
DB      4022 CTCACCGGAAATCTTCAGAGAAACCGACGCTTTTCGATACGGAATCATGCTTTAGA 4081

RESULT 5
X23533
ID      X23533 standard; DNA; 3842 BP.
AC      X23533:
DE      17-JUN-1999 (first entry)
KW      Tomato Xa21 clone TRK2 DNA fragment.
KW      Xa21; receptor kinase-like protein; multigene family; RRR; rice; TRK2;
KW      plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.
OS      Lycopersicon esculentum.
PN      MO9909151-A2.
PD      25-FEB-1999.
PF      17-JUL-1998; U14841.
PR      13-AUG-1997; US-910386.
PA      (REGC ) UNIV CALIFORNIA.
PI      Hubert SH, Richter T, Ronald PC, Song W, Szabo V,
PI      Wang G;
DR      WPI: 99-204431/17.
DR      P-PSDB; W23533.
PT      New RRR polynucleotides and nucleic acid constructs - used for
PT      generating transgenic plants resistant to Xanthomonas
PS      Claim 14; Page 61-62; 67pp; English.
CC      This invention describes a method for conferring disease resistance in
CC      plants. The invention describes the use of novel genes and proteins
CC      belonging to the Oryza longistaminata and Oryza sativa receptor
CC      kinase-like protein (RRK) Xa21 multigene family. Such genes from
CC      cassava, maize and tomato are also described. The genes and proteins can
CC      be used for enhancing resistance to Xanthomonas in a plant, preferably in
CC      rice or tomato.
SQ      Sequence 3842 BP; 984 A; 719 C; 894 G; 1244 T;

Query Match      7.9%; Score 143.4; DB 1; Length 3842;
Best Local Similarity 53.4%; Pred. No. 3.2e-27;
Matches 350; Conservative 0; Mismatches 296; Indels 9; Gaps 2;

QY      800 ttgaaaggttgattgtgtaaggtgtataagggagcctgtgctgactgtgactgtgag 859
DB      2712 TAGGACAGTGAAGCTTTGGAGCAACATTAACAACGGAATTTCACACGAGTCTCTAAGG 2771
QY      860 cagttaaagcgtttaagaagaagacgaacacagcaggtgtgtgagctgcagtttcaacagag 919
DB      2772 CAGTAAAGCAGACTTGCTGTAGAGACGTTTTCAGGAGATT---CAACAGTTGATGCAGAAA 2828

```







PD 25-FEB-1999  
 PR 17-JUL-1998; U14841.  
 PR 13-AUG-1997; US-910386.  
 PA (KRGK ) UNIV CALIFORNIA.  
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 DR WPI: 99-204431/17.  
 PR New RRR polynucleotides and nucleic acid constructs - used for  
 PR generating transgenic plants resistant to Xanthomonas  
 PS Claim 12: Page 58-59; 67pp: English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the *Oryza longistaminata* and *Oryza sativa* receptor  
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
 CC rice or tomato.  
 SQ Sequence 3045 BP; 717 A; 739 C; 744 G; 735 T;

Query Match	7.8%;	Score 142;	DB 1;	Length 3045;
Best Local Similarity	53.9%;	Pred. No. 6.6e-27;		
Matches 336; Conservative	0;	Mismatches 251;	Indels 6;	Gaps 2

OY	800	tttgaagaagtgatgttgtaaggctgataaaggagacgctctgtctatggtcacttgat	859
Db	1988	TTTGCAACAGGAGGATTTGGGTGAGAGATTACAAAGGCTATGGCTATGGATGGCAACGCTTTGGG	2047
OY	860	cagttaaaaggcttaaagaagaacgacacccagctgtgtgagctgcagttccaaacaagg	919
Db	2048	CTGTCAAGAAACTCATGCTATTTTCCACAGGCCAAGGCG--ACGGGGAGTTCACTGACAGAG	2104
OY	920	tggaaatgattagatgatgctgctgcatggaatcttctgcttaagctgttcttcacaga	979
Db	2105	TGGAGACCAATGGCAAAATCAAAATGCAACTTGTGGCGTTGCTAGGCTACTCTCAAG	2164
OY	980	caccaacagagcggctctcttgatatacatatcagctaaigtgaagtgtgtcgttglt	1039
Db	2165	TTGGCGAGCAACGCTCTGCTGTGTACAGATCATGATTAATGGAAAGCCGATGTCTTCG	2234
OY	1040	taagagagcgtgcagacatagaaacccctccctattgtggcaactggaagagatgac	1099
Db	2225	TCCATGAAGGAGACAAGACAGTACTGTGGGTTTATTATTTGGCAACAGAAAGATTGGAG	2284
OY	1100	taggactctgtcaggggcttcttatttgcacatgacattgatacccaagattacatc	1159
Db	2285	TTGGCTCGGCAAGGAGGATGCGCTTCTCCACCATATGTGCATCCGACCATCATACACG	2344
OY	1160	gtgagttaaagaagcgcaatatattatttggacgaagaatttaggcgtgtgtgagtgatt	1219
Db	2345	GGGAACTAATATCAAGCAACGTGTTTGTACATATCTCGATGCTTACGTACGTATGGATT	2404
OY	1220	ttggattagctagagctcatgattacaagaatacccaatgttaca--actcgtgaagg	1276
Db	2405	TCGGATATGGCCGGCTCGTGAATCTGTTATACCAATCAATCAACCGAGCAAGAGCTTTAG	2464
OY	1277	gtaccttgggtacaatagctcccgagtaactctcgactgaaagtcatagaagaagcg	1336
Db	2465	GAAACACGCTTATGTGGCTCCGAGTACTTCCAGTCGGTTATTTCACAACTAAGGGGG	2524
OY	1337	atgtcttgggtatagagattatgtctttaaagctatctactgcagagaagactttgact	1396
Db	2525	ACGTATACAGCTATAGGCGCTGTTTCTTGGAGCTTCTCTAGGGAATAAACCAATCAATC	2584
OY	1397	ttgctcgaccttgcgaacgatgat 1419	
Db	2585	CGACTGAAATTGGCGGACAAATAT 2607	

RESULT	8
T31307	
ID	T31307 standard; cDNA; 1554 BP.
AC	T31307;

DT 15-NOV-1996 (first entry)  
 DE tomato RKK gene clone TRK1.  
 KM Xa21; RKK; disease resistance; Xanthomonas; blight; rice;  
 KW tomato receptor kinase 1; TRK1; transgenic plant; crop protection;  
 RV ss.  
 OS Lycopersicon esculentum.  
 PN MO9623375-A2.  
 PD 25-JUL-1996.  
 PF 17-JAN-1996; U00717.  
 PR 17-JAN-1995; US-373374.  
 PR 07-JUN-1995; US-475891.  
 PR 29-SEP-1995; US-004645.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Ronald PC, Song W, Szabo V, Wang G;  
 DR WPI; 96-354532/35.  
 PT P-PSDB; W03186.  
 PT Plant disease resistance gene Xa21, codes for RKK polypeptide -  
 useful for conferring resistance to Xanthomonas in rice and tomatoes  
 PS Example 4; Page 33; 47pp; English.  
 CC cDNA clones TRK1 (Tj31307) and TRK1 (Tj31308) were isolated from  
 CC a tomato cDNA library by screening with probes generated by PCR  
 CC amplification of tomato cDNA using primers (Tj31301-06) based on the  
 CC rice disease resistance Xa21 gene (Tj31300). They belong to the RKK  
 CC family of disease resistance genes. TRK1 (tomato receptor kinase 1)  
 CC is present as 1 or 2 copies in the tomato genome and 1 copy maps to  
 CC the short arm of chromosome 1 in the proximity of a resistance gene  
 CC to Xanthomonas campestris pv. vesicatoria Rx1. It encodes an RKK  
 CC protein (W03186) that is an important component of plant signal  
 CC transduction pathways leading to a defence response. The gene is  
 CC useful for engineering disease resistance in tomato and other plants  
 CC Sequence 1554 BP; 477 A; 303 C; 348 G; 426 T;

Query Match	7.7%;	Score 140;	DB 1;	Length 1554;
Best Local Similarity	51.8%;	Pred. No. 1.6e-26;		
Matches 423;	Conservative 0;	Mismatches 375;	Indels 18;	Gaps 4;

	Matches	423:	Conservative	0:	Mismatches	375:	Indels	18:	Gaps	4:
QY	793	acacactctggaagagtgagatttgtaagtggtatataaggaagccgttcgtatgctca	852							
Db	715	AACCTCATATAGGGAGAGGTGGACAGGAAATGTTATAGGGCAACATATAGCCCAATATGGTGAT	774							
QY	853	ctttagcacttaaaagcgttaagaagaagcaaacacaggttggctgcgcagttcaca	912							
Db	775	CATGTCCGGGTGAAGAAATTGGGAATAAGCAAAAGGCTACATGATAACGGCCATATCT---	832							
QY	913	acagaagtggaatgatgatcagcgtctgtcgcgcgaacacctctcgtctacgttggttc	972							
Db	832	GCTAACTTAAACATATTAGGGAAGATCAAGCGCATATGATGTAAGAACATCTCGCGT	891							
QY	973	tgcacgacacacaagagcgcgctctctgatatccatcagtcagtaagtgtgtcg	1032							
Db	892	TGTTCCAAACAAGAGATCAACTTGTGATTGTATGATCAATGCTMAATATGGAAAGTTGG--	950							
QY	1033	tcgtgtttaagagagcgctcagccatataagacaccccccttgatttggccaactaggaagag	1092							
Db	950	---GTGAAGTCTTCATGTGGAAAGACGGCGGCAACTCAATATGGAAATAGGCGCTPAAA	1005							
QY	1093	attgcactagatctcgtctaggggcgcttcttttttgcacagccattgatacccaagatt	1152							
Db	1006	ATAGCCATAGAGCTGTGCATAGGGCCCTTTTATTATTATGCAACCAAGATGTGTCCTCATATATA	1065							
QY	1153	atccatcgtgatgtgaaagctgtcaaatatattatttgcacgaagaatttggcgtgtgtga	1212							
Db	1066	ATCCACCGCGCATGTCAAGTCCCAACATATATATGTGAACTCTGAAACTTGAAGCTCATGTT	1125							
QY	1213	ggtatttttgtagttagcgtccacaggtttcaagaagtataccocgt---ttacaacgtc	1269							
Db	1126	GCAGATTTTGGATTAGCCAAAGTACTTTCCTTAACAATGGTACCTCTGATGTGCATGTGTGCA	1185							
QY	1270	gtaagaggctactcgtggtctacatagctcccgagctactctcgacttggaaagtcacag	1329							
Db	1186	ATTCAGAGATCTTATGGCTCATATGCTGCAGAAATATGCAATACAGCTGAAAATATGATAG	1245							



QY 1330 aagaccgattcttcttgatgattatgctcttagagctcttactgacagagagct 1389  
DB 1246 AAAAGCATGTGTATACCTTTGGAGTGTGTTGGAGCTTAAACAGACAAAGCCCA 1305  
QY 1390 ttgatctctgcgccttcgacagatgacatgattatgcttgattggttaaaagc 1449  
DB 1306 GTAGGAATTTTGGAGAAAGAAAGATGACATTGTACAAATGGCGAAACGGAGACAAA 1365  
QY 1450 ctttgaagaagaaaaagttgagatgctgctgacatccctgacctgagaaacattacatt 1509  
DB 1366 TGGAGCAAAAGAGGGGTGTGAAATCTTGG-----ATGAGAGGCTAAAAAAATGTTGCA 1419  
QY 1510 gacacagaagttgacagctattcaagtagcattactctgctgacagggctgcacatg 1569  
DB 1420 ATTGTTAAAGCTATGCAAGATTTTGTGAGCAATGCTTTGTGAGAGTACACATT 1479  
QY 1570 gaggcgcttaagatgctcagagtagctccgaatgctt 1605  
DB 1480 GAGAGGCTTCAATGAGGAAGTAGTCCAAATGCTT 1515

RESULT 9  
ID X23532 standard; DNA; 3293 BP.  
AC X23532;  
DE 17-JUN-1999 (first entry)  
KW Tomato Xa21 clone TRK1 DNA fragment.  
KW Xa21: receptor kinase-like protein; multigene family; RRR; rice; TRK1;  
KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
OS Lycopersicon esculentum.  
PN MO909151-A2.  
PE 25-FEB-1999.  
PF 17-JUN-1998; US-910386.  
PR 13-AUG-1997; US-910386.  
PA (NEGC) UNIV CALIFORNIA.  
PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,  
PI Wang G;  
DR P-PSDB: M93600.  
PT New RRR polynucleotides and nucleic acid constructs - used for  
PT generating transgenic plants resistant to Xanthomonas  
PS Claim 14; Page 59-60; 67pp; English.  
CC This invention describes a method for conferring disease resistance in  
CC plants. The invention describes the use of novel genes and proteins  
CC belonging to the Oryza longistaminata and Oryza sativa receptor  
CC kinase-like protein (RRK) Xa21 multigene family. Such genes from  
CC cassava, maize and tomato are also described. The genes and proteins can  
CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
CC rice or tomato.  
SQ Sequence 3293 BP; 937 A; 634 C; 701 G; 994 T;

Query Match 7.7%; Score 140; DB 1; Length 3293;  
Best Local Similarity 51.8%; Pred. No. 2.2e-26;  
Matches 423; Conservative 0; Mismatches 375; Indels 18; Gaps 4;

QY 793 accactactgaaagaggtgattggtgaaggtgataaagagccttgctgagtgctca 852  
DB 2135 AACGTCATAGGAGAGAGGTGAGGAGGATGATGATAAGGAACTATGCCAAATGGGAT 2194  
QY 853 ctgtgacgattaaaggcttaagaagaagcaaccaggtggtgagctgacgtttcaa 912  
DB 2195 CATGTCCGGGTGAAGAATAATGGGAATAGCAAAAGCTCACATGATTAACGGCCTATCT 2252  
QY 913 acagaagtggaatgattgattgctgctgctgacatctcttcgctcactggtttc 972  
DB 2252 GCTGAACCTTAACATTAAGGAAATGACGATAGTACATGTGAGACGCTGCGCTTT 2311  
QY 973 tgcattgacacacagagcggctctctgatatcatcatcagtgatgagtggtgcg 1032  
DB 2312 TGTTCMAACAAGAAAGTCAACTTGTCTAGTTATAGATACATCAATTAATGGAAGCTTGA-- 2370  
QY 1033 tcgtgtttaagagagcgctcagcattcagaccccttcctgattggtccaactagaagag 1092

DB 2370 ----GTGAAGCTTCATGGAAGAAGCGGCGCAATCCAAATGGAAATAGGCTTAAA 2425  
QY 1093 attgactaggatcgtcgttaagggtctcttatttcataagacattgatacccaagatt 1152  
DB 2426 ATAGCCATGAAAGCTGCCAAGGGCCTTCTTATTTGCACACAGATTCCTCCCATGATA 2485  
QY 1153 atccatgtagtataaaagctgcgaatatattatgagcagaagaagttagggctgtgta 1212  
DB 2486 ATCCACCGCGATGTCATCCCAACAAATATATTGTTGAACCTTGAACTTGAAAGCTCATGTT 2545  
QY 1213 gttatttttggttagctagctcatgattacaaagatacccatg---ttacaactgct 1269  
DB 2546 GCAGATTTTGGATTATGCAATACTTTCGTACAAATGTTACCTTGATGCAATGTCGCA 2605  
QY 1270 gtaagggttaccttggtgctacatagctcccgagatccctcagctgagaaagtcacag 1329  
DB 2606 ATTGCAGATCTTATGCTTACATTGCTCCAGAAATATCATACAGCTGAAATATGATGAG 2665  
QY 1330 aagaccgattccttggtatgattatgctcttagagctcattactgacagagagct 1389  
DB 2666 AAAAGCATGTGTATGAGCTTTGGAGTGTGTTGGAGCTTAAACAGAGCAAGGCCA 2725  
QY 1390 ttgatcttgcgccttcgacagatgattatgcttcttgattggttaaaagc 1449  
DB 2726 GTAGGAATTTTGGAGAAAGAAAGATGACATTGTACAAATGGCGAAACGAGACAAA 2785  
QY 1450 ctttgaagaagaaaaagttgagatgctgctgacatccctgacctgagaaacattacatt 1509  
DB 2786 TGGACCAAGAAAGAGGGGTGTGAAATCTTGG-----ATGAGAGGCTTAAAAAAATGTTGCA 2839  
QY 1510 gacacagaagttgacagcttattcaagtagcattactctgctgacagggctgcacatg 1569  
DB 2840 ATTGTTGAAGCTATGCAAGTATTTTGTACCAATGCTTTGTGAGAGTACAGCATT 2899  
QY 1570 gaggcgcttaagatgctcagagtagctccgaatgctt 1605  
DB 2900 GAGAGGCTTCAATGAGGAAGTAGTCCAAATGCTT 2935

RESULT 10  
ID T62124 standard; cDNA to mRNA; 3176 BP.  
AC T62124;  
DE 17-JUN-1997 (first entry)  
KW Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.  
KW Plant; morphogenesis; regulation; short; stem; alteration;  
KW inflorescence; extraneous; gene; expression; transformation;  
KW increase; control; form; length; ds.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT cds 51..2981  
FT /tag= a  
FT /note= "plant morphogenesis regulatory protein"  
PN J09056382-A.  
PD 04-MAR-1997.  
PE 24-AUG-1995; 216187.  
PR 24-AUG-1995; JP-216187.  
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.  
PA (CHIK-) ZH CHIKUO KANKYO SANGYO GIJITSU KENKYU.  
DR WPI: 97-206629/19.  
DR P-PEDB: W13408.  
PT DNA encoding plant morphogenesis regulatory protein - useful to  
PT yield plants with short stems or altered inflorescence  
PS Claim 1; Pages 6-10; 17pp; Japanese.  
CC The present sequence encodes an Arabidopsis thaliana plant  
CC morphogenesis regulatory protein (MRP), which can be used to yield  
CC a plant with, e.g. short stems or altered inflorescence. The MRP  
CC acts on a plant at a specific site for a specific period, and can  
CC therefore be used to regulate extraneous gene expression in a  
CC plant. The MRP's cDNA or genomic DNA can be used to transform a  
CC plant to increase its MRP expression, and therefore control the  
CC form (particularly stem length) of the plant.



SQ Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match 7.6%; Score 138.4; DB 1; Length 3176;  
Best Local Similarity 56.5%; Pred. No. 5,6e-26;  
Matches 278; Conservative 0; Mismatches 211; Indels 3; Gaps 1;

QY 904 cagttccaaacagaggtgtaaatgattagcagctgctgcatcgaatcttcctgctcta 963  
|||||  
DB 2112 CAGTTGAAACGAACTCGAGATGCTAACTACATCAACAGCAAGAACTCTGTGAGCTTA 2171  
QY 964 cgtggtttctcagtcagcaacaagcgcctctgtatataccatacagctatgta 1023  
DB 2172 CAGGCTTATTCCCTCTCTCACTTTGGGAGTCTTCTGTTATGATATTGGAATAATGCT 2231  
QY 1024 agtgttcgctcgtgtttaaagaagcgtcagccatcagaacctcccttgattgccaact 1083  
DB 2232 AGCCTTGAGATCTTCTTATGCCCCCTACGAAGAAAAGACT--CTTGATTGGGACACA 2288  
QY 1084 aggaagagattgcactagatctgctagaggcttcttatttcagbacattgtgat 1143  
DB 2289 CGGCTTAAGTACATATGCTGACACAAAGTTTAGCTTATCTACACATAGCTGTAGT 2348  
QY 1144 cccaagatatccatcgtatgttaaagctgcaatatataatttggacgaagattgag 1203  
DB 2349 CCAAGGATCATCTCAGAGACGAGAGTGTCCCAACATTCTTGGACAAAAGACTTAGAG 2408  
QY 1204 gctgtgttggtgatttttggttagctagctcattgataaagaatacccatgttaca 1263  
DB 2409 GCTCGTTTACACATTTTGGAAATGCGAAAGCTGTGTGTCAAAAGTCAATCTTCA 2468  
QY 1264 actgcgttaagagggtaccttggtctacatagctcccgagtaacctcgcagctggaagtica 1323  
DB 2469 ACTTAGTGTATGGCGACGATAGGTTACATAGACCCGAAATATGCTGCGACTTCAAGGCTC 2528  
QY 1324 tcagagaagacagatcctttgttataggatatagtctcttagagctcattactgacaag 1383  
DB 2529 ACTGAGAAATCCGATGCTACAGTATGGAATAGTCCCTTAGAGCTTTAACCCGAAGG 2588  
QY 1384 agagctttgat 1395  
|||||  
DB 2589 AAGCCGCTTGAT 2600

RESULT 11  
V22790  
ID V22790 standard; CDNA: 2640 BP.  
AC V22790;  
DT 23-JUL-1998 (first entry)  
DE Class I S-receptor kinase (SRK) gene nucleotide sequence.  
KW Class I S-receptor kinase; SRK; Brassica campestris s45; ds.  
OS Brassica campestris.  
FH Key Location/Qualifiers  
FT CDS 45..2585  
FT /tag- a  
PN J10099076-A.  
PD 21-APR-1998.  
PR 26-SEP-1996; 254945.  
PA (SALS-) SAISHU JITSUYO GIJUNSU KENKYUSHO KK.  
DR WPI; 98-289869/26.  
DR P-PSDB; W56307.  
PT S-locus glycoprotein gene - and s receptor kinase gene of class I  
PT type and class II type to produce recombinant plants  
PS Disclosure; Fig 2; 20pp; Japanese.  
CC The present sequence represents a class I S-receptor kinase (SRK)  
CC gene. The sequence was isolated from Brassica campestris s45 CDNA library.  
SQ Sequence 2640 BP; 800 A; 491 C; 645 G; 704 T;

Query Match 6.7%; Score 121.2; DB 1; Length 2640;  
Best Local Similarity 53.7%; Pred. No. 1.3e-21;  
Matches 321; Conservative 0; Mismatches 268; Indels 9; Gaps 3;

QY 791 gtacatacttgaagaggtgatttgtaagtgatataaggagacctgtctatgct 850  
|||||  
DB 1630 GTAAACAACCTCGGACAGGTGGTTGGATTGTTTAAAGGAAAGATTACTTACCGGCGC 1689  
QY 851 cacttgaagcagttaaagagcttaagaagaacagacacccagtggtgtagctgcaattc 910  
|||||  
DB 1690 AAGAAATCCGTGTAATAAGGCTATCAGAGACCTCAGTTCMAOGACTATG--AGTTTA 1746  
QY 911 aaacagaagtggaatattgatacagctgctgcatcgaatcttcctgctcagctggt 970  
DB 1747 TGAATGAGGTGACATTAATAGCTAGAGCTTCACACATATAACCTTGTCAAAATTCTTGACT 1806  
QY 971 tctcagtcagccacaagaagcgcctctgttatataccatacagctatgtaagattgtg 1030  
DB 1807 GTTGCAATTGAGCAGATGAGAAAGATGCTATATATGAGTATTGGAATAATTTAAAGCTTCG 1866  
QY 1031 cgtcgtgtttaagaagagcgtcagccatcagaacctcccttgattgccaactagaaga 1090  
DB 1867 ATTCTTATCTCTCG--GAAAAACCCAAAGGTCATAGCTAATTTGGAAGAGAGATTGCG 1923  
QY 1091 ggaattgcactagatctgctagagggtcttcttatttcagacattgatacccaaga 1150  
DB 1924 ACATTAACCAATGGTGTGCTCGAGGCTTTTATCTTATCTTATCAAGACACACGTTTAAAGA 1983  
QY 1151 ttatccatcgtgtgttaaagctgcaatatataatttggacgaagaatttgaagctgttg 1210  
DB 1984 TAATCCACAGAGATTTAAAGTAAGTAACATTTGCTTGATATAAATATGATCCAAAGA 2043  
QY 1211 taagtatttgggttagctagctcagctcagatataaagaataccatgt--acaactg 1267  
DB 2044 TCTCGATTTTGGGATGGCCAGGATATTGCAAGGACGAGACGGAAGCTAACACAAATGA 2103  
QY 1268 ctgtaaggggtacctgggtcgtacatagctcccgagtaacctctcagctggaagttcataag 1327  
DB 2104 AGGTCGTGGAACCTTACGCTACATGTCCCGAGAGTACCGATGATGGGATATTTCTCGG 2163  
QY 1328 agaagacagatgcttctgtgtataggatattcgtcttaagctcattactgacagag 1385  
|||||  
DB 2164 AAAATACAGATGTTTCAAGTTGTTGAGTCAATAGTCTTGAATTTGTACTGGAAGAG 2221

RESULT 12  
Q32648  
ID Q32648 standard; DNA: 2833 BP.  
AC Q32648;  
DT 04-MAY-1993 (first entry)  
DE S-receptor kinase gene SRK6.  
KW S-receptor; kinase; SRK; pistil; anther; stigmatic; SI; response;  
KW self-incompatibility; primary; transmitter; pollen recognition;  
KW signal; S-locus; binding domain; transmembrane; protein kinase;  
KW glycoprotein; Slg; SRK6; SRK2; ZMPK1; ss.  
OS Brassica oleracea.  
FH Key Location/Qualifiers  
FT CDS 1..2574  
FT /tag- a  
FT signal\_peptide 1..97  
FT /tag- b  
FT msc\_rna 1340..1399  
FT /tag- C  
FT /note= "Transmembrane domain"  
PN EP-519869-A.  
PD 23-DEC-1992.  
PR 10-JUN-1992; 810449.  
PR 19-JUN-1991; US-717331.  
PA (CIRBA ) CIRBA GENIV AG.  
PI Nasrallah UB, Nasrallah ME, Stein J;  
DR WPI; 92-426051/52.  
DR P-PSDB; R29814.  
PT New Brassica gene for S-receptor kinase, used to impart self  
PT incompatibility to plants - has extracellular domain homologous  
PT to S-locus glyco-protein gene  
PS Claim 8; Page 51-52; 81pp; English.



CC The sequences given in Q32648-49 represent two S-receptor kinase (SRK)  
 CC genes. These genes are transcribed specifically in pistils and  
 CC anthers and are involved in the stigmatic self-incompatibility (SI)  
 CC response of Brassica. The protein products of these genes are thought  
 CC to be the primary transmitters of the pollen recognition signal in the  
 CC SI response. The SRK proteins comprise an S-locus binding domain, a  
 CC transmembrane domain and a protein kinase domain. The S-locus binding  
 CC domain is pref. substantially homologous to an S-locus glycoprotein  
 CC (SLG) gene. The kinase domain has a degree of homology with the  
 CC kinase domain of SRK6, SRK2, and ZMPK1.  
 CC Sequence 2833 BP; 857 A; 541 C; 646 G; 789 T;

Query Match 6.4%; Score 115.4; DB 1; Length 2833;  
 Best Local Similarity 53.1%; Pred. No. 4.1e-20;  
 Matches 317; Conservative 0; Mismatches 271; Indels 9; Gaps 3;

OY 791 gtacatactggaagaggtgattggttaaggtgtatataagagagccttgatgct 850  
 DB 1592 GTAAACAACCTGGACAGAGTGTATGTTATGTTAAAGGAAGATTACTGACGGGA 1651  
 OY 851 cactgtgcaagttaaagagccttaagaagaacacagcggtggtgagctgagtttc 910  
 DB 1652 AAGAATTCAGTAAAGGCTATCAAGAGCTCAGTCAAGGAGCTGATG--AGTTTA 1708  
 OY 911 aacagaggtggaagtagatgcatgctgtgcatcgaaatctctgctgctagctgctt 970  
 DB 1709 TGAAATGAGTGTACCTAATTTGCGAGGCTTCACGATATTAACCTGTTAACTTCTG 1768  
 OY 971 tctgcatgacacacacagagcgctctctgtatatacatagagagagaggtgtg 1030  
 DB 1769 GTTGCATTTGAAGAGATGGAAGATGTTGATATGATGATTTGGAAATTTAAGCCTTG 1828  
 OY 1031 cgtcgtgttaagaagagcgctcagcatcaagaacctccctggtgagctgcaactggaaga 1090  
 DB 1829 ATTCCTTACT--CTATGTGTAACCCGAAAGGCTTACCTAATTTGATGAGATTCG 1885  
 OY 1091 gattgacactagatcctgaaggggtctcttatttcagtagacattgatacccaaga 1150  
 DB 1886 ACATTACCAATGCTGTGCTCGAGGCTTTTATATCTTCACTCAAGACTCAAGGTTTAA 1945  
 OY 1151 ttatccatcgtgtaagtaaaagctgcaaatatataattgagaagaattgagctgtg 1210  
 DB 1946 TATCCACAGAGATTGGAATGAATACATTTTCTTGACAAAATATGATCCCAAGA 2005  
 OY 1211 taggtgatttgggttagctagctcagcatgattacaagatacccatgtt--acaactg 1267  
 DB 2006 TCTCGATTGTTGGATTGCGCAGGATTTTGAAGAGGACGAAACGAACTTAACACAAATGA 2065  
 OY 1268 ctgtaaggggtacacttggtctacatagctcccgagctacactcagctggaagatcatag 1327  
 DB 2066 AGTGTGTCGACATACGCGCTACATGTCGCCGATACGCAATGATATGGAATTTCTG 2125  
 OY 1328 agaagacagatgcttgggtatggtatgctcttagagctcattacgtagacaga 1384  
 DB 2126 AAAAATCAGATGTTTTCAGTTTGGAGTCATAGTCTTGAATTTGTAGTGAAGA 2182

RESULT 13  
 ID T13425 standard; cDNA; 2685 BP.  
 AC T13425;  
 DT 10-JUN-1996 (first entry)  
 DE SRK-A10 cDNA.  
 KW S locus receptor kinase; SRK-A10; S locus glycoprotein; SLG-A10;  
 KW self-incompatibility; hybrid seed; transgenic plant; rapeseed;  
 KW Brassica napus oleifera line W1.  
 OS Brassica napus oleifera line W1.  
 FT Key Location/Qualifiers  
 FT cds 36..980  
 FT misc\_difference 948  
 FT /\*tag= a  
 FT /\*tag= b

FT /note= "there is a single base-pair deletion at  
 FT position 948"  
 PN CA2123751-A.  
 PD 12-SEP-1995.  
 PF 17-MAY-1994; 123751.  
 PR 11-MAR-1994; US-208909.  
 PR 11-MAY-1994; US-242104.  
 PA (RORH/) RORHSTEIN S J.  
 PI Glavin TL, Goring DR, Rothstein SJ, Schafer U;  
 DR WPI: 96-180298/19.  
 PT DNA mols. encoding S locus glycoprotein or receptor kinase -  
 PT expression vectors contg. the DNA can be used to produce  
 PT self-incompatible plants, which ensure hybrid progeny and increase  
 PT yield  
 PS Disclosure: Fig 2; 3pp; English.  
 CC cDNA clone SRK-A10 (T13425) codes for an S locus receptor kinase  
 CC associated with the A10 self-incompatibility allele of Brassica  
 CC napus. It was isolated from genomic DNA of rapeseed line W1 by  
 CC PCR amplification. The gene contains a 1 bp deletion resulting in  
 CC premature termination of translation and produ. of truncated SRK-A10  
 CC protein. Expression vectors contg. a corrected SRK-A10 gene (see  
 CC T13423) are used to produce self-incompatible plants able to  
 CC produce hybrid seed in improved yields.  
 CC Sequence 2685 BP; 816 A; 512 C; 639 G; 718 T;

Query Match 6.3%; Score 114.8; DB 1; Length 2685;  
 Best Local Similarity 53.0%; Pred. No. 5.7e-20;  
 Matches 317; Conservative 0; Mismatches 272; Indels 9; Gaps 3;

OY 792 taccatactggaagaggtgattggttaaggtgtataagagagccttgatgctc 851  
 DB 1624 TAACAATAATTGACACAGGCGCTTTTGGTATTTTTCACAAAGGAGATTAATCTGACGGCA 1683  
 OY 852 actgtagaagttaaaggcttaagaagaacacacaggtgtgagctgagttcca 911  
 DB 1684 AGAAATCGCGGTAAAGAGCTATCAAAAGCTGACGTTCAAGGCTGTATG--AGTTAT 1740  
 OY 912 aacagaggtggaatgattagatgctgtgcatggaatctctgctgctagctgtt 971  
 DB 1741 GAATGAGGTGACATTAATGCGAGGCTTCAACATGTAATCTTGCAAAATCTTGCGCTG 1800  
 OY 972 ctgacagacacacagagcgctctctgtatatacatagatgcatggaaggtgtgc 1031  
 DB 1801 TTGCATTGACGCGATGAGAAGATGCTGATATATGATGTTTGAATAATTAAAGCTCGA 1860  
 OY 1032 gtcgtgttaagaagagcgctcagcatcagaacctccctgattggtccaactagaagag 1091  
 DB 1861 TTCTTATCTCTTCTG--GAAAAAATCTGAAGGCTTAAGCTAATTTGGAAGAGAGATTGGA 1917  
 OY 1092 gattgacatagatctgctgaaggggtctcttatttcagtagacattgatacccaagat 1151  
 DB 1918 CATTCACATGCTGTGCTCGAGGCTTTTATATCTTCAAGACTCCCGCTTATGAT 1977  
 OY 1152 tatccatcgtgtaagtaaaagctgcaaatatataattgagaagaattgagctgtgt 1211  
 DB 1978 AATCCACAGAGATTGGAAGTAAGTAACATTTTGTCTGATGAAATATGTCGCCAAAGAT 2037  
 OY 1212 aggtgatttgggttagctagctcagcatgattacaagagatacccatgtt--acaactgc 1268  
 DB 2038 CTCGGATTTTGAATGCGCAGGATATTGGAAGAGCGAGAGGAGGATTCACATGTA 2097  
 OY 1269 tftaaggggtacacttggtgctacatagctccgagtagcctcagcagtggaagatcatcaga 1328  
 DB 2098 GGTGCTGGAACTTACGGCTACATGTCGCCAGAGTACGAAATGGGTGGGATTTCTCGGA 2157  
 OY 1329 gaagacagatgcttgggtatggtatggtcctcattagagctcattactcagtagacaga 1386  
 DB 2158 AAATACGATGTTTTCAGTTTGGAGTCATGCTTGAATTTATTTACTGGAAGAAGA 2215

RESULT 14  
 ID T13423



ID T13423 standard; cDNA; 2686 BP.  
AC T13423;  
DT 10-JUN-1996 (first entry)  
DE SRK-A10 gene mutant N948.  
KW S locus receptor kinase; SRK-A10; S locus glycoprotein; SLG-A10;  
KW self-incompatibility; hybrid seed; transgenic plant; rapeseed;  
KW canola; oilseed; vector; ds.  
OS Synthetic.  
FH Key  
FH mutation  
FH location/Qualifiers  
FT 948  
FT /tag= a  
FT /note= "a, c or g may be inserted at position 948"  
PN CA123751-A.  
PD 12-SEP-1995.  
PE 17-MAY-1994; 123751.  
PR 11-MAR-1994; US-208909.  
PR 11-MAY-1994; US-242104.  
PA (ROMH/) ROTHSTEIN S U.  
PI Glavin TL, Goring DR, Rothstein SJ, Schaefer U;  
PI WPI; 96-180298/19.  
PT DNA mols. encoding S locus glycoprotein or receptor kinase -  
PT expression vectors contg. the DNA can be used to produce  
PT self-incompatible plants, which ensure hybrid progeny and increase  
PT yield  
PS Claim 8; Fig 3; 39pp; English.  
CC cDNA clone SRK-A10 (T13423) codes for an S locus receptor kinase  
CC associated with the A10 self-incompatibility allele of Brassica  
CC napus. The gene contains a 1 bp deletion resulting in premature  
CC termination of translation and prodn. of truncated SRK-A10  
CC protein. A 'corrected' gene (T13423) has a, c or g inserted at  
CC position 948 (T13423). Incorporation of the corrected gene into  
CC an expression vector allows the breeding of self-incompatible  
CC plants able to produce hybrid seed in improved yields.  
SQ Sequence 2686 BP; 816 A; 513 C; 638 G; 718 T;

Query Match 6.3%; Score 114.8; DB 1; Length 2686;  
Best Local Similarity 53.0%; Pred. No. 5.7e-20;  
Matches 317; Conservative 0; Mismatches 272; Indels 9; Gaps 3;

QY 792 taccactactggaagaggtggaattgtaagtgatataaggaagccctgctgcatgctc 851  
DB 1625 TACACAAATTTGGACAGAGGTGTTTGATTTGTTACAAAGGGGAAATTAATCTTGACGGCA 1684  
QY 852 actctgtagcagtaaaagcttaaaagaagaacacacaggtggtgagctgcagttca 911  
DB 1685 AGAAATCGGGTAAAGGCTATCAAGACGACAGTCAAGGGGTTGAG---AGTTTAT 1741  
QY 912 aacagaagtggaatgataagcagctgctgcatcgaaactcttcgctctagctgttt 971  
DB 1742 GAATGAGTGACATTAATCGAGGCTTCAACATGTAATCTTGCAAAATCTTGCGTG 1801  
QY 972 ctgcataagacacacagagcgctctctgtatatacatcagctgaatgaaagctgttc 1031  
DB 1802 TTGCATTTGACGCGATGAAGAAGTCTGATATGATGATTTTGAAAAATTTAAGCTCGA 1861  
QY 1032 gtccgtttaagaagcgctcagccatcagaacctccccttgattgccaactaggagaag 1091  
DB 1862 TTCTTATCTCTCG---GAAAAACTCGAAGGCTTAAGCTTAATTTGGAAGAGATTCGA 1918  
QY 1092 gattgactagatctgctgaggggcttctatttgcatagacattgcatcccaagat 1151  
DB 1919 CATTACCAATGCTGTGCTCGAGGGCTTTATATCTTCATCAAGACTCCCTTAGGAT 1978  
QY 1152 tatccatcgtatgtaaaagcgtcaaatatattatgacgaaagattggagctgttct 1211  
DB 1979 AATCCACAAGATTGAAAGTAAGTAACATTTTCTGTGATAGAAATATGCTCCAAAGAT 2038  
QY 1212 aggttttttggttagctagctcattgattacaagaagctaccatgtt---acaactgc 1268  
DB 2039 CTGGGATTTTGGATGCGCAGGATATTGAAAGAGCAGAGGAAAGCTTAACCAATGA 2098  
QY 1269 tctaagggttaacctgggtgctaagctcccgagtaacctctgactggaagatcatcaga 1328

DB 2099 GGTGTCGGAACCTTACCGCTACATGTCCCGAGAGTACGCATGCGTGGGATATTCTCGCA 2158  
QY 1329 gaagaccagatgctctgtgattagtgatgctccttagagctcatattaccggaagaga 1386  
DB 2159 AAAATCAGATGTTTTCAGTTTGGAGTCATGATGTTTGAATATTAATCTGGGAAGAGA 2216  
RESULT 15  
T62125  
ID T62125 standard; DNA; 9295 BP.  
AC T62125;  
DT 10-JUN-1997 (first entry)  
DE Arabidopsis thaliana plant morphogenesis regulatory protein DNA.  
KW Plant; morphogenesis; regulation; short; stem; alteration;  
KW inflorescence; extraneous; gene; expression; transformation;  
KW increase; control; form; length; ds.  
OS Arabidopsis thaliana.  
FH Key  
FH location/Qualifiers  
FH 1803.1881  
FH /tag= a  
FH /tag= a  
FH /tag= a  
FH /tag= b  
FH /tag= c  
FH /tag= c  
FH /tag= d  
FH /tag= d  
FH /tag= e  
FH /tag= e  
FH /tag= f  
FH /tag= f  
FH /tag= g  
FH /tag= g  
FH /tag= h  
FH /tag= h  
FH /tag= i  
FH /tag= i  
FH /tag= j  
FH /tag= j  
FH /tag= k  
FH /tag= k  
FH /tag= l  
FH /tag= l  
FH /tag= m  
FH /tag= m  
FH /tag= n  
FH /tag= n  
FH /tag= o  
FH /tag= o  
FH /tag= p  
FH /tag= p  
FH /tag= q  
FH /tag= q  
FH /tag= r  
FH /tag= r  
FH /tag= s  
FH /tag= s  
FH /tag= t  
FH /tag= t  
FH /tag= u  
FH /tag= u  
FH /tag= v  
FH /tag= v  
FH /tag= w  
FH /tag= w  
FH /tag= x  
FH /tag= x  
FH /tag= y  
FH /tag= y  
FH /tag= z  
FH /tag= z  
FH /tag= aa  
FH /tag= aa



FT Intron 4282..4349  
FT /tag- ab  
FT Exon 4350..4421  
FT /tag- ac  
FT Intron 4422..4508  
FT /tag- ad  
FT Exon 4509..4580  
FT /tag- ae  
FT Intron 4581..4706  
FT /tag- af  
FT Exon 4707..4778  
FT /tag- ag  
FT Intron 4779..4860  
FT /tag- ah  
FT Exon 4861..4932  
FT /tag- ai  
FT Intron 4933..5018  
FT /tag- aj  
FT Exon 5019..5090  
FT /tag- ak  
FT Intron 5091..5176  
FT /tag- al  
FT Exon 5177..5248  
FT /tag- am  
FT Intron 5249..5412  
FT /tag- an  
FT Exon 5413..5481  
FT /tag- ao  
FT Intron 5482..5576  
FT /tag- ap  
FT Exon 5577..5648  
FT /tag- aq  
FT Intron 5649..5726  
FT /tag- ar  
FT Exon 5727..5800  
FT /tag- as  
FT Intron 5801..5882  
FT /tag- at  
FT Exon 5883..6011  
FT /tag- au  
FT Exon 6096..6443  
FT /tag- av  
FT Intron 6012..6095  
FT /tag- aw  
FT Intron 6444..6519  
FT /tag- ax  
FT Exon 6520..5412  
FT /tag- ay  
FT Intron 6891..6974  
FT /tag- az  
FT Exon 6975..7328  
FT /tag- ba  
PN J09056382-A.  
PD 04-MAR-1997.  
PE 24-AUG-1995; 216187.  
PR 24-AUG-1995; JP-216187.  
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.  
PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.  
DR WPI: 97-206629/19.  
PT DNA encoding plant morphogenesis regulatory protein - useful to  
PT yield plants with short stems or altered inflorescence  
PS Claim 6; Pages 12-15; 17pp; Japanese.  
CC The present sequence encodes an Arabidopsis thaliana plant  
CC morphogenesis regulatory protein (MRP), which can be used to yield  
CC a plant with, e.g. short stems or altered inflorescence. The MRP  
CC acts on a plant at a specific site for a specific period, and can  
CC therefore be used to regulate extraneous gene expression in a  
CC plant. The MRP's cDNA or genomic DNA can be used to transform a  
CC plant to increase its MRP expression, and therefore control the  
CC form (particularly stem length) of the plant.  
SQ Sequence 9295 BP; 2873 A; 1581 C; 1602 G; 3239 T;

Query Match 6.2%; Score 112.6; DB 1; Length 9295;  
Best Local Similarity 54.5%; Pred No. 3.6e-19;  
Matches 226; Conservative 0; Mismatches 189; Indels 0; Gaps 0;  
QY 1067 cccttgatggccaactaggaagagatgacactagatctgtagggcttcttatt 1126  
DB 6538 CTCTTGATGGGACACACGGCCTTAAGATAGCATATGTGCACACAAAGTTAGCTTATC 6597  
QY 1127 tgcataaccaattgatacccaagattatccatcctgtagtaaaagctgcaataattat 1186  
DB 6598 TACACATGACTGTAGTCCAGGATCATTCACAGAGACGTGAAGTGTCCAACTTCCT 6657  
QY 1187 tggacgaagaattggaaggtctgtgtagtgatttgggtacgaaggtcactgattaca 1246  
DB 6658 TGGACAAAGACTTAGAGGCTCGTTTGACAGATTTTGGAAATACGAAAGCTTGCTGCT 6717  
QY 1247 aagatacccatgtacaactcgtgtaaggggtaccttgggtacataagctccgaatacc 1306  
DB 6718 CAAGTCACTACTTCAACTTACGTATGAGGCGACATAGGTTACATAGACCCGAGTATG 6777  
QY 1307 tctcgactggaagatcatcagaagaacggatgcttgggtatgggtatgatagccttag 1366  
DB 6778 CTCGCACTTACGGCTCCTCAGAGAAATCCGATGCTTACAGTTATGCAATAGTCTTCTTG 6837  
QY 1367 agctcaactactggaagaagaagcttggatccttgcgccttgcgaacgaatgata 1426  
DB 6838 AGTTGTTAACCAGAAAGACCGGTGATGACCAATCCATCCACATCGCTTGTGT 6897  
QY 1427 tgtgttgatctgggttaaaagccttltgaagaagaagaagtggagatgctggt 1481  
DB 6898 CTTCCTTGCTTACTCTCTCAGCTGCTGCTGTTAGGTTCAAGTCCGTAATCTTGT 6952

Search completed: June 23, 2000, 22:49:30  
Job time: 40366 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:41:28 ; Search time 274.21 Seconds  
(without alignments)  
860.373 Million cell updates/sec

Title: US-09-180-798-2  
Perfect score: 1815  
Sequence: 1 gacaaatcacatgaaatcat.....ttaaaaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/5C.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/5D.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/6.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/PCTUS.COMB.seq:\*  
7: /cgn2\_6/ptodata/2/1na/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	7.7	1554	US-08-587-680A-24	Sequence 24, Appl
2	117	6.4	2571	US-07-717-331F-9	Sequence 9, Appl1
3	117	6.4	2833	US-07-717-331F-1	Sequence 1, Appl1
4	110.4	6.1	2749	US-08-265-628-1	Sequence 1, Appl1
5	99.6	5.5	966	US-08-447-185-2	Sequence 2, Appl1
6	99.6	5.5	2443	US-08-447-185-3	Sequence 3, Appl1
7	92	5.1	2749	US-07-717-331F-4	Sequence 4, Appl1
8	87.6	4.8	5733	US-08-473-533A-1	Sequence 1, Appl1
9	58.6	3.2	3921	US-08-587-680A-3	Sequence 3, Appl1
10	58.6	3.2	5992	US-08-475-891A-3	Sequence 3, Appl1
11	57.2	3.2	4517	PCT-US93-06251-83	Sequence 83, Appl
12	57.2	3.0	2075	US-08-238-163-3	Sequence 3, Appl1
13	54.8	3.0	6236	US-08-475-891A-1	Sequence 1, Appl1
14	53.8	3.0	6256	US-08-567-375-1	Sequence 1, Appl1
15	53.8	3.0	6256	US-08-587-680A-1	Sequence 1, Appl1
16	53.8	3.0	2880	US-08-666-271-4	Sequence 4, Appl1
17	52.6	2.9	3905	US-08-666-271-1	Sequence 1, Appl1
18	52.6	2.9	2647	PCT-US93-06251-77	Sequence 77, Appl
19	52.4	2.8	2890	US-07-928-464-1	Sequence 1, Appl1
20	51.4	2.8	2890	PCT-US93-07347-1	Sequence 1, Appl1
21	51.4	2.8	3033	US-08-003-311B-1	Sequence 1, Appl1
22	51.4	2.8	3033	US-08-261-432-1	Sequence 1, Appl1
23	50.8	2.8	1804	US-08-306-691B-40	Sequence 40, Appl
24	50.8	2.8	1804	PCT-US93-06251-82	Sequence 82, Appl
25	49.8	2.7	1038	US-08-238-163-1	Sequence 1, Appl1
26	49.8	2.7	831	US-08-567-375-15	Sequence 15, Appl

28	47.6	2.6	1203	US-08-602-010A-7	Sequence 7, Appl1
29	47.6	2.6	1203	US-08-680-726A-7	Sequence 7, Appl1
30	47.6	2.6	5495	US-08-602-010A-1	Sequence 1, Appl1
31	47.6	2.6	5495	US-08-602-010A-2	Sequence 2, Appl1
32	47.6	2.6	5495	US-08-680-726A-1	Sequence 1, Appl1
33	47.6	2.6	5495	US-08-680-726A-2	Sequence 2, Appl1
34	47.6	2.6	10592	US-08-680-726A-51	Sequence 51, Appl
35	47.6	2.6	10592	US-08-680-726A-52	Sequence 52, Appl
36	46.2	2.5	3590	US-08-587-689-1	Sequence 1, Appl1
37	46.2	2.5	3590	PCT-US96-09193-1	Sequence 1, Appl1
38	46	2.5	4626	US-08-306-691B-22	Sequence 22, Appl
39	46	2.5	4626	PCT-US93-06251-27	Sequence 27, Appl
40	45	2.5	7218	US-08-232-463-14	Sequence 14, Appl
41	44.6	2.5	3389	US-08-395-580-1	Sequence 1, Appl1
42	44.6	2.5	3426	US-08-205-018-1	Sequence 1, Appl1
43	43.8	2.4	1602	US-07-820-011A-1	Sequence 1, Appl1
44	43.8	2.4	1602	PCT-US93-00445-1	Sequence 1, Appl1
45	42.2	2.3	10350	US-09-047-125-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1  
US-08-587-680A-24  
; Sequence 24, Application US/08587680A  
; Patent No. 5977434  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Confering  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/587, 680A  
; FILING DATE: 17-JAN-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373, 375  
; FILING DATE: 17-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475, 891  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004, 645  
; FILING DATE: 29-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/567, 375  
; FILING DATE: 04-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34, 774  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1554 base pairs



```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA (partial)
US-08-587-680A-24

```

```

Query Match          7.7%: Score 140; DB 4; Length 1554;
Best Local Similarity 51.8%; Pred. No. 4.2e-29;
Matches 423; Conservative 0; Mismatches 375; Indels 18; Gaps 4;

```

```

QY 793 accaactctggaagaggtggaatttgtaaggtgataaagagacgcttctgcatgctca 852
DB 715 AACGTCATAGGAGAGGAGTGAGCAAGGATAGTATATAAGGAAGTAATGACCAATGGAT 774
QY 853 ctgttagcagtttaaaaggcttaaaagaaacgaaacagagtggtgagctgcaattcaa 912
DB 775 CATGTCGGGGTGAAGAAATTTGGGATATAGCAAAAGGCTCATGATGATACGGGCTATCT 832
QY 913 acagaggtggaagatgattagatgctgtgcatgaaatctctgctgctacggtgttc 972
DB 832 GCTGACTTAAACACTTATGGGAGATCAGGCAATAGGTAATTTGAGAGACTGCTGCGCTTT 891
QY 973 tgcattgacacacacagagcagcttctgtatatacatatagcttaattggaagtgttcg 1032
DB 892 TGTTCACAAAGAGAGTCAACTTGTAGTTATGATGATCATGCTAATATGGAAGCTTAG-- 950
QY 1033 tctgttttaaaagagcgtcagccatcagacccctctgtattggtccaactagaagag 1092
DB 950 ---GTGAAGTCTTCAATGAGGAGAGACGGCGCACTCCATGGAAGGAAATGAGCTTAA 1005
QY 1093 attgacacagatcgtcgaaggggtcttcttatttgcacatgacatgtgacccaagatt 1152
DB 1006 ATAGCCATAGAAAGCTGCCAAGGCGCTTCTTATTTGGACCAAGATGCTGCCCTCATGATA 1065
QY 1153 atccatcgtatgataaaagctgcaataatattattggaagaaattggaagctgtgtta 1212
DB 1066 ATCCACCCGAGTGCACATCCACAAATATTTGTAATCTGGAACCTGGAAGCTGACATGTT 1125
QY 1213 ggttgatttggttagcaggtcctatgataaagatgatacaggttacacgt---ttacac 1269
DB 1126 GCAGATTGTGATTAGCCAAAGTACTTGTATCAATAGGTACTCTGTGATGATGTCTGCA 1185
QY 1270 gtaaggggttaaccttggtacatagctccagctacccctcagactggaagatcacagag 1329
DB 1186 ATGAGAGATCTTATGGCTATGCTCAGAAATATGATGATGATGATGATGATGATGATGAT 1245
QY 1330 aagacagatgcttctgttataaggtatctctcttaagagctcattactggaagagagct 1389
DB 1246 AAAACCGATGCTATAGCTTTGGAGTGGTGTGTTGGAGCTTATTAACAGGACGAAAGCCA 1305
QY 1390 ttgtatctgctcgcctggtcgaagcagatgatttattggtgttggttggttaaaagc 1449
DB 1306 GTAGCAAAATTTTGGAGAAAGAAAGATGACATTTGATCAATGGGCAAAACGAAACAAA 1365
QY 1450 ctttgaagagaaaagctggaagatgctgctgacccctgacccctggaagaaatatact 1509
DB 1366 TGGACAAAGAGAGGCTGTAATAATCTTGG-----ATGAGAGGCTAATAAATGTTGCA 1419
QY 1510 gacacagaagtgtgagcagctattcaagtaagcattactctgtaccagaggttgcacagt 1569
DB 1420 ATTGTGAAGCTATCAAGTATTTTGTGACAAAGCTTGTGTGTAAGAGTACAGCATT 1479
QY 1570 gagcgcttaagatgtcagaggttadtcgaagtctt 1605
DB 1480 GAGAGGCTTACATAGGGAAGTACTCCAAATGCTT 1515

```

```

RESULT 2
US-07-717-331F-9
; Sequence 9, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:

```

```

; APPLICANT: June Nasrallah, Michael Nasrallah, and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-07-717-331F-9

```

```

Query Match          6.4%: Score 117; DB 1; Length 2571;
Best Local Similarity 53.3%; Pred. No. 1.1e-22;
Matches 318; Conservative 0; Mismatches 270; Indels 9; Gaps 3;

```

```

QY 791 gtaacactctggaagaggtgatttgtaaggtgataaagagacgcttctgcatgctc 850
DB 1592 GTAACAACTCGAACAAGTGGTTTGGTATTTACAAAGGAAGATTACTTACGGGA 1651
QY 851 cactgtgacatttaagagcttaagaagaagaaacaccccggtggtgagctgacttcc 910
DB 1652 AAGAAATGCAATAAAGGCTATCAAAAGACGTCAAGTCAAGGAGACTGATG--AGTTTA 1708
QY 911 aaacagaggtggaatgatttaagcatgctgtgcatcgaatctctgctcagctggt 970
DB 1709 TGATAGAGTGACATAATTTGCGAGGCTTCAGCANATTAACCTGTTCAAGTTTGCT 1768
QY 971 tctgcatgacacacagagcggctcttctatatacatcacatagcttaattggaagtgtg 1030
DB 1769 GTTGATTGAAGAGATGAGAGATGATGATATGATATTTGGAATAATTAAGCCTTG 1828
QY 1031 cgtggtttaaagagaggtcagcatagaaaccccttgattggtgccaactggaga 1090
DB 1829 ATTCTATCTCTTGG---GTAAACCCGAAAGGCTTAAGCTAATTAATGGAATGAGATTCG 1885
QY 1091 ggaattgacatgagatccttaaggggtcttctatttgcataagcattgtatcccaaga 1150
DB 1886 ACATTACCAATGTTGTGTCGAGGCTTTATATCTTCATCAACAGCTACACGTTTGA 1945
QY 1151 ttatccatctgtaattaaagctcaaatatattattggaagaaattggaagctgttg 1210
DB 1946 TAATCCACAGAGATTGGAAGTAAGTAACATTTGCTTGACAAAATAATGATCCCAAGA 2005
QY 1211 taggtgatttggttagtagctcagatgattacaagatgataccacagctt---acaactg 1267
DB 2006 TCTCGGATTTGGAGTGGCCAGGATATTTGAAGGAGCAAGAACGACTAACACAAATGA 2065

```



QY 1031 cgtcgtgtttaagagagcgtccagccatcagaacctccccctgatctggccaactagaaga 1099  
||| | | | | | | | | | | | | | |

```

; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; STRAIN: oleifera
;

```

INDIVIDUAL ISOLATE: WI



```

; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: S-locus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2574
; PUBLICATION INFORMATION:
; AUTHORS: GORING, DAPHNE
; AUTHORS: ROTHSTEIN, STEVEN J.
; TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A
; TITLE: FUNCTIONAL SERINE/THREONINE KINASE
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749
; US-08-265-628-1

```

```

Query Match 6.1%; Score 110.4; DB 2; Length 2749;
Best Local Similarity 52.5%; Pred. No. 7.7e-21;
Matches 315; Conservative 0; Mismatches 276; Indels 9; Gaps 3;

```

```

QY 789 tagtaccatacttgaaagaggtgattgtgaaggtgtataaagggagccttcgtatgg 848
DB 1593 ttgttaacaaactcgacaaaggtggttcggtattgtttacaaaggtatgattactgatgg 1652
QY 849 ctcccttgtagcagttaaaaggttaaaagaaacagacacaggtgtagtgcagtt 908
DB 1653 gcagaaatttcgggttaaaaggtatcaaaaacgtcggttcagagagctgctg---agtt 1709
QY 909 tcaaacagaggtggaatgattagcagtgctgtagcatcgaatctctcgtctacgtgg 968
DB 1710 tatgaatgagctgagattgattcgagagcttcacatataaacttgcctccgaattcttgg 1769
QY 969 ttctcgacagcacaaacagagcggtctctgtatcatcatagcttaagaaagtg 1028
DB 1770 ctgttgcatgtagcagacgaaagatgctggtatgattttagaaaattttacgct 1829
QY 1029 tgcgtcgtgttaaagagagcgctcagcagaaaccccttgatttggcacaactagaa 1088
DB 1830 cgattcttatctcttcg---gaaatataacgaagctctacgtaattgaagacagatt 1886
QY 1089 gagggttgcactagatctgttaggggtcttcttatttcagtacatctgtatcccaa 1148
DB 1887 caacattacaaatggtgttgcagagacgtttttatcttcacaaagctcagcgttttg 1946
QY 1149 gattaccatcgtatgtaaaagctgcacatatatttgagcagaagatttgaagctgt 1208
DB 1947 gataatccacagagatataagaaagtaagtaacattttgcttgatataaaatgacacaa 2006
QY 1209 tgtaggtgatttgggttagctaggtctcagttacaaagatacccaagtt---acaac 1265
DB 2007 gatctcggaattttggagtgcccaaatctttgcaagggagacgaactgaactcaacag 2066
QY 1266 tgcgttaagggttaccctgggtctacatagctccagtagtacctcgcagctgaaagtc 1325
DB 2067 gaagtggtcggaaacttccgctacatgctccggaatagcgaatgaaatggggtattctc 2126
QY 1326 agagaagacagatgcttctgttctatggaatagcctccttagagctcatctcgtacag 1385
DB 2127 gaaaaaatacagatgttttcaagttttgagatcattgttgaattgttagtgaaaaag 2186

```

# RESULT 5

```

; Sequence 2, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATHOGE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051

```

```

; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,078
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman Mr., Michael L.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1600
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-447-185-2

```

```

Query Match 5.5%; Score 99.6; DB 1; Length 966;
Best Local Similarity 51.5%; Pred. No. 4.1e-18;
Matches 340; Conservative 0; Mismatches 299; Indels 21; Gaps 4;

```

```

QY 752 ttctctcgcgaatttcgaagtcgacaggaacttttagtaca-----tacttgaa 805
DB 86 ttctctttagatgttgagaggaagcaactatattttgacacagttttatatttgac 145
QY 806 gagggtgatttggtaaggtgtatagaagcgtctgtcgtatgcttactgttagcagtt 865
DB 146 atggctgttttggagaggtttacaaaggtgtttgcgtatgagcaaaagtgccctga 205
QY 866 aaggtcttaagaagaaacgaacacacaggtgtgtagctgcagtttcaacagaaggttgaa 925
DB 206 aaagcgttaccgacgagctcctcacaagattttgaag---agttgaaacgaatgttga 262
QY 926 tgattagcattgctgtcattcgaatccttctcgtctcagtggttcttcgatacaca 985
DB 263 ctctctcatttttgcaagacatccgcatctggttcattgatatgagattcttgatgaaga 322
QY 986 cagagcgtctctgtatcatcatatgcttaaggaatgtgtcgtcgtgttgaag 1045
DB 323 atgacatatttcttatttataatgacatgagaaatgggaactcacaagacattttatg 382
QY 1046 agcgtcagccatcgaacccctcgtattggtccaaactaggaaggaattgcataagat 1105
DB 383 gatcagatcttaccacaaatgacatgagctggagacagagctggaatgacatgagg 442
QY 1106 ctgctagggtgtcttctatttgcattgcatgcatgtgtgtccaaagattatcatcgtgatg 1165
DB 443 cagccagaggtc-----tacacttacttcttacttaagcaatttatcactgcatg 493
QY 1166 taaagcgtcaaatatattatgagcaagaatttgaagctgtgttagtatttgggt 1225
DB 494 tcaacttataaactatgttcttgaatgaaattttgcccataaaatactagattttgaa 553
QY 1226 tagctagctcattgattacaaggt---accatgttacaactgctgtatagggtaact 1282
DB 554 tatccaaagaaagagctgagcttgcatacaacccatcttaccacagatgagtaagagc 613
QY 1283 tgggtatagctcccgagtaactctcgaactgaaaggtcatcagaagacagatgct 1342

```



Db 386 ATGAGATGATCTAATTTATAATACATGAGAGATGGAAACCTCAAGACATTTGTATG 445

QY 1046 aagcgtacgccatcagaacctccccttgatttggccaacttaggaagagattgcaactggat 1105

Db 446 GATCAGATCTACCCACAAATGACCAATGACACTGGAGCAGAGAGCTGGAATATGCAATGGGG 505

QY 1106 ctgcctaggggsgcttccttctatttgcatagcacatgtgatacccaagatattccalcgtgac 1165

Db 506 CACCAGAGAGTCTC-----TACACTACCTTACTACTAGAGCAATATATACATCGTAGTG 556

QY 1166 taaagctgcacaatatattatttgcagcaagaatttggagctgtgtgcattgatttggat 1225

Db 557 TCAAGCTATTAACAAATATGCTGTGATGGAATTTGTGCCAAAAATACGATTTTGGAA 616

QY 1226 taccctgagctcatgagttcacaagat---accatgttacaactgcgtgtaagggtaccct 1282

Db 617 TATCCAGAAAGAGGACTGACCTGTGATCAAAACCCATCTTTAGCAGCTGATGAAGAGCACTC 676

QY 1283 tgggctacatagctcccgagtaacctctcgcactggaaagatcatcagagaagacagatgct 1342

Db 677 TCGGCTACATATGACCTGAAATATTTATTAAGAGGACGACTCAGTGAATAATCTGATGTT 736

QY 1343 ttggattaggtatbagcctcttagagccattactcagtcagacagagcctttgatctgtc 1402

Db 737 ATTCTTGGGTGTTGTTTATTGGAAGTCTTGTGTAGGCTGCGCATGATGTCATCTC 796

RESULT 7

US-07-717-331F-4

Sequence 4, Application US/07717331F

Patent No. 5484905

GENERAL INFORMATION:

APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua

APPLICANT: Stein

TITLE OF INVENTION: A Receptor Protein Kinase Gene

TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Yahwak & Associates

STREET: 25 Skytop Drive

CITY: Trumbull

STATE: Connecticut

COUNTRY: USA

ZIP: 06611

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: Macintosh

OPERATING SYSTEM: MS-DOS

SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/717.331F

FILING DATE: June 19th 1991

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: George M. Yahwak

REGISTRATION NUMBER: 26,824

TELECOMMUNICATION INFORMATION:

TELEPHONE: (203)268-1951

TELEFAX: (203)268-1951

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2749 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-07-717-331F-4

Query Match 5.1%; Score 92; DB 1; Length 2749;

Best Local Similarity 50.5%; Prid. No. 8.7e-16;

Matches 305; Conservative 0; Mismatches 290; Indels 9; Gaps 3



QY	785	cttttgtagccacactctgtgaagaaggggtgagcttggaaggtgtgataaaggacacgccttgctg	844
Db	1604	CTGATTTTAAACAAGCTCGGAAAAGGTGGCTTTGGTGTCTTTTACAAGGGAAGCTTAGTTG	16633
QY	845	atggtcactctgtgagagttaaagaaggtcttaaagaagaacgaacccaggtyggtgacgtgc	904
Db	1664	ACGGGCAAGAAATTCGAGTCGAAGAGACTATACGGAATATGCACTCAAGGTACCGATAGAT	17233
QY	905	agtttcaaaagaggtggaagaatgatagcatggtcgtgcatcgtgaatcttcgtgtac	964
Db	1724	TCAATGAACGAAGTTAGGCTTAATCAAAAGCTTACGCCACAAT--AATCTGTGCGACCTTC	17600
QY	965	gtgtgtctctcatgagacaacacagagcggtctcttgatataccatctgctaaagaa	10244
Db	1781	TGGCGCTGTGTGTTTATGAGCGCGAGAAAGATCTTAATTACGAGTACTTGGAAATCTAA	18400
QY	1025	gtgtgtcgtctgtttaaagagagcgtcagccatcagaacctccccttgattggtccaacta	10848
Db	1841	GCCCTCGATTCATCTCTTTTGA--TGAAACCAAGAGCTGTATGTTAAATGGCAAAATGA	18977
QY	1085	ggaagagatctgacactagagatctgcgcgaaggsggtcttctatttgcatgacactgtgac	11444
Db	1898	GATTGATATTAATCAATGATGATATGGCCGAGGCTTCTATCTTCAACCAAGTTAACGCT	19557
QY	1145	ccaagattatccatcgtgatagttaaagctgycaaatatattatttggacgaagaattgag	12044
Db	1958	TTAGAAATCAATCCACAGGAGATTGAAAGCAAGCAATGTCTGCTTGATAAAGATATGACTC	20117
QY	1205	ctgtgtcgaagtgatttgggttcgcctgacgctcgtatgaataagaatataccactgtt---	12648
Db	2018	CAAAAATTTTCAGACTTTGGATGGCTTAGGGATCTTTTGACGGATGAGCGGAAGCTTGACA	20777
QY	1262	caactcgtctgaaggggtacaccttgggtctacaagctcccgagtaacctctgcactgtgaagt	13212
Db	2078	CGAAGAAAGTGTGTCGGAATTTATGGCTACATGCTCCAGAAATATGCGATGAACGGACAT	21337
QY	1322	catagaagaagaccgatgctcttgggttatggattatgctcttagagctcatctagac	13878
Db	2138	TCTCAATGAAGATCGAAGTGTTCAGTTTGGGGCTTGCTTGTGAATTAATTAAGTGCA	21977
QY	1382	agag 1385	
Db	2198	AGAG 2201	
<p>RESULT 8</p> <p>US-08-473-553A-1</p> <p>: Sequence 1, Application US/08473553A</p> <p>: Patent No. 5859338</p> <p>: GENERAL INFORMATION:</p> <p>: APPLICANT: Meyerowitz, Elliot M.</p> <p>: APPLICANT: Clark, Steven E.</p> <p>: APPLICANT: Williams, Robert W.</p> <p>: TITLE OF INVENTION: Plant Clavatal Nucleic Acids,</p> <p>: TITLE OF INVENTION: Transformed Plants, and Proteins</p> <p>: NUMBER OF SEQUENCES: 11</p> <p>: CORRESPONDENCE ADDRESS:</p> <p>: ADDRESSEE: Flehr, Honbach, Test, Albritton &amp; Herbert</p> <p>: STREET: Four Embarcadero Center, Suite 3400</p> <p>: CITY: San Francisco</p> <p>: STATE: California</p> <p>: COUNTRY: United States</p> <p>: ZIP: 94111-4187</p> <p>: COMPUTER READABLE FORM:</p> <p>: MEDIUM TYPE: Floppy disk</p> <p>: COMPUTER: IBM PC compatible</p> <p>: OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>: SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>: CURRENT APPLICATION DATA:</p> <p>: APPLICATION NUMBER: US/08/473,553A</p> <p>: FILING DATE: 06-JUN-1995</p> <p>: CLASSIFICATION: 800</p> <p>: ATTORNEY/AGENT INFORMATION:</p>			

```

NAME: SILVA, Robin M.
REGISTRATION NUMBER: 38, 304
REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5733 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 2434..5037
FEATURE:
NAME/KEY: CDS
LOCATION: 5117..5467
US-08-473-553A-1

Query Match          4.8%; Score 87.6; DB 3; Length 5733;
Best Local Similarity 51.3%; Pred. No. 2,1e-14;
Matches 232; Conservative 0; Mismatches 214; Indels 6; Gaps 1;

OY   793  acctacttggaagagtgatattgtagtgtataagaagcgcttgtatgctca 852
      |||||
DB   4534  AACGTATGCGTAAAGCCGCAGTGATTTGTCTACCGTGCATCATGCCAACACGTA 4593

OY   853  ctgtacgagttaaaaggcttaagaagaacgaaccaggltggcgtgcagttccaa 912
      |||||
DB   4594  GACGTCCCGATTAAACGACTCCTGGCCGTGGACCGGAGAGACGATCAATGATTACAG 4653

OY   913  accgaagtggaatgttatgcatgctgtgatgaaatcttcgctcacgtgttttc 972
      |||||
DB   4654  GCGGAGATTCAAACCTTTGGGGGAATTCGCCACCGETCACATAGTAGACTTCTTGCGTTTC 4713

OY   973  tgcatagcaccaacagagcgcttcctgtlatatccatatcatgctaagtgaagtgttcg 1032
      |||||
DB   4714  GTACCGAACAAGATTCGAATCTCTCTCTTATGATGATACATGCAATGGAAGCCTTGGAA 4773

OY   1033  tcgtgtttaagagagcgctaacgccatcgaacctccccttgatttggccaactagaagaag 1092
      |||||
DB   4774  GAGGTTTT-----GCATGATCTAAAGGTGGTATCTTCANTGGAGAGAACAATATA 4827

OY   1093  atgcataagatcgtcgtgagggcttcattattgtcatacgtacatcttgcatacgaatt 1152
      |||||
DB   4828  GTAGCCCTGGAACCTGCACAAAGGCTTGCTATCTTCACCATGATTTTACCATGATATC 4887

OY   1153  atccatcgtatgttaaagctgcaaatatbatatbtatgtagcagaagaattggagctgttgta 1212
      |||||
DB   4888  TTGCATGTGAAGATGTTAAGTCCAAATTAACATCTTTTGGACCTGATTTTGAAGCCCATGT 4947

OY   1213  ggtgatttgggttagctagctcattgatta 1244
      |||||
DB   4948  GCTGATTTGGGCTCTAAAGTTCTTAATTGA 4979

RESULT           9
US-08-567-375-3
Sequence 3, Application US/08567375
Patent No. 5952485
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
APPLICANT: Szabo, Veronique
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Disease Resistance in Plants
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:

```



```

: ADDRESS: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/567,375
: FILING DATE: 04-DEC-1995
: CLASSIFICATION: 800
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 60/004,645
: FILING DATE: 29-SEP-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/475,891
: FILING DATE: 07-JUN-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/373,375
: FILING DATE: 17-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34,774
: REFERENCE/DOCKET NUMBER: 023070-058930
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3921 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(1..2676, 3520..3918)
: OTHER INFORMATION: /product= "Xa-21"
:
: US-08-567-375-3
:
: Query Match 3.2%; Score 58.6; DB 4; Length 3921;
: Best Local Similarity 50.1%; Pred. No. 1.6e-06;
: Matches 172; Conservative 0; Mismatches 169; Indels 2; Gaps 1;
:
: QY 86 aattctatagatcgtaacagatataataattacatgacgttcacatgacct 145
: DB 1126 ACTTCGCTTAGTTTCTTCTGCACTTGATTAATGAATGACACAGGAAGCATTCGGAAGGAT 1185
:
: QY 146 acccttgacaatatggggtctctatgacatgagctttacagcaataacataagtgac 205
: DB 1186 ATTGGCAATCTTATTTGGCTTACAAACATC--TCTATCTCTGCAACAACAATTTCAGAGG 1243
:
: QY 206 caattcctagatccttggaatcctgacaatttggtgagcttgagcctatataatgata 265
: DB 1244 CTCCTTCATCATCGTTGGGCGAGGCTTAAAACTTAGGCATTTACCGCTTACGAAACA 1303
:
: QY 266 gcttccttgacacctatacggagacacacattaggaagccttaagcctaagcttcttcgctc 325
: DB 1304 ACTTGAGCGGTTGCATCCCGTTGGCCATAGGAATCTTACTGAACTTAATATCTTTACTGC 1363
:
: QY 326 tcaacaacaacagcctctcgttccaaatccaaatgacactgactgataatatacaactcttc 385
: DB 1364 TCGGCACCAACAATTCAGGTGTGATACCATACACATCTCAAAACCTCAAACTTGT 1423
:
: QY 386 aagtcctcgattatcaaaacatcgctatcaggaccagtaacc 428
: DB 1424 TGTCATTAGGCTTCAACTAATTAACCTTAGTGCTCAATAC 1466
```

```

: RESULT 10
: US-08-587-680A-3
: Sequence 3, Application US/08587680A
: Patent No. 5977434
: GENERAL INFORMATION:
: APPLICANT: Ronald, Pamela C.
: APPLICANT: Wang, Guo-Liang
: APPLICANT: Song, Wen-Yuang
: APPLICANT: Szabo, Veronique
: TITLE OF INVENTION: Procedures and Materials for Confering
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/587,680A
: FILING DATE: 17-JAN-1996
: CLASSIFICATION: 800
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/373,375
: FILING DATE: 17-JAN-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/475,891
: FILING DATE: 07-JUN-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 60/004,645
: FILING DATE: 29-SEP-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/567,375
: FILING DATE: 04-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34,774
: REFERENCE/DOCKET NUMBER: 023070-058940US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3921 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(1..2676, 3520..3918)
: OTHER INFORMATION: /product= "Xa-21"
:
: US-08-587-680A-3
:
: Query Match 3.2%; Score 58.6; DB 4; Length 3921;
: Best Local Similarity 50.1%; Pred. No. 1.6e-06;
: Matches 172; Conservative 0; Mismatches 169; Indels 2; Gaps 1;
:
: QY 86 aattctatagatcgtaacagatataataattacatgacgttcacatgacct 145
: DB 1126 ACTTCGCTTAGTTTCTTCTGCACTTGATTAATGAATGACACAGGAAGCATTCGGAAGGAT 1185
:
: QY 146 acccttgacaatatggggtctctatgacatgagcctttacagcaataacataagtgac 205
: DB 1186 ATTGGCAATCTTATTTGGCTTACAAACATC--TCTATCTCTGCAACAACAATTTCAGAGG 1243
```



Query Match 3.2% Score 58.6 DB 3 Length 5992

Y 206 caattctagtatcttgggaatctgaacaatttggtagcttggaacctatactgaata 265

Db 1244 CTCTTCATCATGATGTTGGGAGAGCGCTTAAAAACTTGAAGCATTTACTCGGCTCAGAAACA 1308

Y 266 gctctcttgaccatactccggaacacataagaaagcttacaagctaaagctatctgcgc 325

Db 1304 ACTTGACGGGTTGCATCCCTGTGGCCATAGGAATCTTACTGAACCTTATATCTTACG 1363

Y 326 tcaacaacaagagctctcgtgcacaattccaatgacatgacatgactaataataactctc 385

Db 1364 TCGGCACCAACAATTCAGTGTGGTTCATCATTACACACTCTCAAACTCAGCAAACTGT 1423

Y 386 aagtcctgattatcaacaacatcggcgtacatcaggaccagtaacc 428

Db 1424 TGTCAATAGGCGCTTCACTAAATAACCTTGTGTGTCATACC 1466

RESULT 11

US-08-475-891A-3

Sequence 3, Application US/08475891A

Patent No. 5859339

GENERAL INFORMATION:

APPLICANT: Ronald, Pamela C.

APPLICANT: Wang, Guo-liang

APPLICANT: Song, Wen-xiang

TITLE OF INVENTION: Procedures and Materials for Confering

TITLE OF INVENTION: Disease Resistance in Plants

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475,891A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/373,375

FILING DATE: 17-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 02370-058910US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5992 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: join(512..3149, 3993..4393)

OTHER INFORMATION: /product= "RRK-B"

OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease

OTHER INFORMATION: resistance gene RRK-B from rice (Oryza

OTHER INFORMATION: sativa)"

US-08-475-891A-3

	Best Local Similarity 50.1%: Pred. No. 2e-06;	Matches 172;	Conservative 0;	Mismatches 169;	Indels 2;	Gaps 1;
QY	86	aattctatagaaatcgtaacagatataatataatataatcattgacggttaactgattgctt	145			
Db	1637	ACTTCGGTTTGGTTTCTTCTTACCTTAATGTGAATGAATCAGAGAAAGCAATTCGAAAGGAT	1696			
QY	146	accttgcaaatatggtggtcttaagacttggagctttacagcaataaataaagtggac	205			
Db	1697	ATTGGCAATTTTATTGCGCTTTACAAATC--TCTATCTCTGCAACAAATTTTCAAGGGT	1754			
QY	206	caattccatgatactcttgggaactgcgaacaatttggtagcttggacctatacatgata	265			
Db	1755	CTTTCCATCATGGTTGGGCGAGGCTTTAAAACTTAGGCATTTTACCTGCTTACGAAATA	1814			
QY	266	gtctcttgacctataaccgagacaattaggagaagcttaacaggtctaagaattcttgcgc	325			
Db	1815	ACTTGACCGGTTTGATCCCGTTGGCCATAGGAATACTTACGAACCTTAATATTTCCTCC	1874			
QY	326	tcaacaacacagcctctcttgcaccaattccaaatgcatcgaactaataatbacaaactcc	385			
Db	1875	TGCGACCAACAATTCATGTGTGGATGGATACCATACACCTCTGTAACACTTCACAAACTTGT	1934			
QY	386	aagtcctggaattatcaacaacatcggtctatcaggacccagttacc	428			
Db	1935	TGTCATTAGGCTTTTCACTAATAATACCTTAGTGCGCAATTC	1977			

```

1      RESULT 12
2      PCT-US93-06251-83
3      ; Sequence 83: Application PC/TUS9306251
4      ; GENERAL INFORMATION:
5      ; APPLICANT: Wikstrom, Eric and Rife, Jason P.
6      ; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
7      ; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
8      ; NUMBER OF SEQUENCES: 93
9      ; CORRESPONDENCE ADDRESS:
10     ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
11     ; STREET: 400 Garden City Plaza
12     ; CITY: Garden City
13     ; STATE: NY
14     ; COUNTRY: USA
15     ; ZIP: 11530
16     ; COMPUTER READABLE FORM:
17     ; MEDIUM TYPE: Floppy disk
18     ; COMPUTER: IBM PC compatible
19     ; OPERATING SYSTEM: PC-DOS/MS-DOS
20     ; SOFTWARE: PatentIn Release #1.0, Version #1.25
21     ; CURRENT APPLICATION DATA:
22     ; APPLICATION NUMBER: PCT/US93/06251
23     ; FILING DATE: 19930630
24     ; CLASSIFICATION:
25     ; ATTORNEY/AGENT INFORMATION:
26     ; NAME: Digiglio, Frank S.
27     ; REGISTRATION NUMBER: 31,346
28     ; REFERENCE/DOCKET NUMBER: 8586
29     ; TELECOMMUNICATION INFORMATION:
30     ; TELEPHONE: 516-742-4343
31     ; TELEFAX: 516-742-4366
32     ; TELEX: 230 901 SANS UR
33     ; INFORMATION FOR SEQ ID NO: 83:
34     ; SEQUENCE CHARACTERISTICS:
35     ; LENGTH: 4517 base pairs
36     ; TYPE: nucleic acid
37     ; STRANDEDNESS: double
38     ; TOPOLOGY: linear
39     ; MOLECULE TYPE: DNA (genomic)
40     ; PCT-US93-06251-83
41
42     Query Match 3.28; Score 57.2; DB 6; Length 4517;
43     Best Local Similarity 53.08; Pred. No. 4.1e-06;
44     Matches 122; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```



QY 1147 aagttctccatcgtgtgtgtgaagaagctgcgaatatatttttgacgaagaattttagct 1206

Db 1378 AACATATTTCACGAGATCTTGGGCTGCTAATATTCTTGTGGGAAAACTGTGTGC 1437

QY 1207 gtctgaagtatttgggtgttagctaaagctcatgattacaaagataccatglttaact 1266

Db 1438 AAAATACACAGACTTGGTTAGCAAGSTTAATGAAGCAATGAMTACACAGACAGCA 1497

QY 1267 gctctaaagggttaccttgggctcatagctcccgagttacctctgactcggaaagtcata 1326

Db 1498 GGTGTCAAAATTTCCAAATCAAAATGGAAGAGCTCCGGAAGCTGCACTGTATGGTGGTTACA 1557

QY 1337 gagaaagccagatgctcttgggttagatcatgctcttagagctcaatc 1376

Db 1558 ATAAAGCTGATGTCTGTGCTATTGGAAATTTGCAACAGAACTGTATAC 1607

RESULT 13

```

US-08-238-163-3
/ Sequence 3, Application US/08238163
/ Patent No. 5569830
/
/ GENERAL INFORMATION:
/
/ APPLICANT: BENNETT, Alan
/ APPLICANT: LABAYVITCH, John M.
/ APPLICANT: POWELL, Ann
/ APPLICANT: STOTZ, Henrik
/ TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
/ TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
/ NUMBER OF SEQUENCES: 24
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend Hourie and Crew
/ STREET: Stewart Street Tower, One Market Plaza
/ CITY: San Francisco
/ STATE: California
/
/ COUNTRY: US
/ ZIP: 94105-1493
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/238,163
/ FILING DATE: 03-MAY-1994
/ CLASSIFICATION: 800
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 2307E-540
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 543-9600
/ TELEFAX: (415) 543-5043
/
/ INFORMATION FOR SEQ ID NO: 3:
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2075 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: DNA (genomic)
/
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 421..1401
/
/
/ US-08-238-163-3

```

Query Match	3.08;	Score 54.8;	DB 1;	Length 2075;
Best Local Similarity	51.7%;	Pred. No. 1.2e-05;		
Matches 125; Conservative	0;	Mismatches 117;	Indels 0;	Gaps 0;

OY 190 ataacataagtgaccaattccatgtaaccttggaaatgcacaatttgytgagcttg 249  
| | | | | | | | | | | | | | | | | |  
Db 721 ACTAATCTCACCGGAACAATTCCACCTGCAATTGGGAAGTCACAAATCTCAAAATGTTA 780

OY 250 gaacttaacttaaatagctctctcggacactaacgggaacastingaaacttaacag 309  
 Db 781 AGGCTACGCTTCACCTACCTTACACAGTCCCACTCCCTGAATTCCTTAGTCAGGTGAAGAT 840  
 OY 310 ctaagaattcttgcgtctcaacaacaagaagctctctcgttcacattacatgtcact 369  
 Db 841 TTGACGTTGTGTGAGTTGAATTAACATCAATTTACCGGAACAAATCCCTTTCCTCTCT 900  
 OY 370 aatattaacaactctcaagtccttgattatcaaaacaatcggctatcaggaccgtacg 429  
 Db 901 CAGCTTCGCAATTTGCTAGCGATGACTTGTAGATGTAACAACCTCACCGGAACAATACCG 960  
 OY 430 ga 431  
 Db 961 GA 962

RESULT 14

```

US-08-475-891A-1
Sequence 1, Application US/08475891A
Patent No. 5859339
GENERAL INFORMATION:
APPLICANT: Ronald, Pamela C.
APPLICANT: Wang, Guo-Liang
APPLICANT: Song, Wen-Yuang
TITLE OF INVENTION: Procedures and Materials for Conferring
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,891A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,375
FILING DATE: 17-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 02370-05891005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(1648..4383, 5178..5513)
OTHER INFORMATION: /product= "RRK-P"
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
sativa)"
US-08-475-891A-1

```

Query Match 3.08; Score 53.8; DB 3; Length 6256;







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:05:52 ; Search time 6198.48 Seconds  
(without alignments)  
1186.839 Million cell updates/sec

Title: US-09-180-798-2  
Perfect score: 1815  
Sequence: 1 gacaaataccatgtaatat.....ttaaaaaaaaaaaaaaaaaa 1815

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : ESR:\*  
1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: em\_est10:\*  
11: em\_est11:\*  
12: em\_est12:\*  
13: em\_est13:\*  
14: em\_est14:\*  
15: em\_est15:\*  
16: em\_est16:\*  
17: em\_est17:\*  
18: em\_est18:\*  
19: em\_est19:\*  
20: gb\_est1:\*  
21: gb\_est2:\*  
22: gb\_est3:\*  
23: gb\_est4:\*  
24: gb\_est5:\*  
25: gb\_est6:\*  
26: gb\_est7:\*  
27: gb\_est8:\*  
28: gb\_est9:\*  
29: gb\_est10:\*  
30: gb\_est11:\*  
31: gb\_est12:\*  
32: gb\_est13:\*  
33: gb\_est14:\*  
34: gb\_est15:\*  
35: gb\_est16:\*  
36: gb\_est17:\*  
37: gb\_est18:\*  
38: gb\_est19:\*  
39: gb\_est20:\*  
40: gb\_est21:\*  
41: gb\_est22:\*  
42: gb\_est23:\*  
43: gb\_est24:\*  
44: gb\_est25:\*

45: gb\_est26:\*  
46: gb\_est27:\*  
47: gb\_est28:\*  
48: gb\_est29:\*  
49: gb\_est30:\*  
50: gb\_est31:\*  
51: gb\_est32:\*  
52: em\_est20:\*  
53: em\_est21:\*  
54: em\_est22:\*  
55: em\_est23:\*  
56: em\_est24:\*  
57: em\_est25:\*  
58: em\_est26:\*  
59: gb\_est33:\*  
60: gb\_est34:\*  
61: gb\_est35:\*  
62: gb\_est36:\*  
63: gb\_est37:\*  
64: gb\_est38:\*  
65: em\_est27:\*  
66: em\_est28:\*  
67: em\_est29:\*  
68: em\_est30:\*  
69: gb\_est39:\*  
70: gb\_est40:\*  
71: gb\_est41:\*  
72: gb\_est42:\*  
73: gb\_est43:\*  
74: gb\_est44:\*  
75: em\_est31:\*  
76: em\_est32:\*  
77: em\_est33:\*  
78: em\_est34:\*  
79: gb\_est45:\*  
80: gb\_est46:\*  
81: gb\_est47:\*  
82: gb\_gss1:\*  
83: gb\_gss2:\*  
84: gb\_gss3:\*  
85: gb\_gss4:\*  
86: em\_gss1:\*  
87: em\_gss2:\*  
88: em\_gss3:\*  
89: em\_gss4:\*  
90: gb\_gss5:\*  
91: gb\_gss6:\*  
92: gb\_gss7:\*  
93: gb\_gss8:\*  
94: gb\_gss9:\*  
95: em\_gss5:\*  
96: em\_gss6:\*  
97: em\_gss7:\*  
98: em\_gss8:\*  
99: em\_gss9:\*  
100: em\_gss10:\*  
101: em\_gss11:\*  
102: gb\_gss10:\*  
103: gb\_gss11:\*  
104: em\_gss12:\*  
105: gb\_gss12:\*  
106: gb\_gss13:\*  
107: gb\_gss14:\*  
108: gb\_gss15:\*  
109: gb\_gss16:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



No.	Score	Match	Length	DB	ID	Description
1	398.4	22.0	656	38	AA738547	AA738547 SBRK5 So
2	365.6	20.1	991	38	AA738544	AA738544 SBRK2 So
3	361	19.9	532	79	AM233982	AM233982 sf32805.Y
4	352.6	19.4	555	51	AI728030	AI728030 BNLGH1962
5	349	19.2	1435	38	AA738545	AA738545 SBRK3 So
6	332	18.3	538	62	AI900110	AI900110 sc01a04.Y
7	279.8	15.4	555	62	AI896277	AI896277 EST265720
8	274.6	15.1	597	51	AI727837	AI727837 BNLGH1920
9	268.2	14.8	655	46	AA738546	AA738546 SBRK4 So
10	250.4	13.8	703	51	AI729440	AI729440 BNLGH1133
11	220.2	12.3	471	47	AI486453	AI486453 EST244774
12	220.6	12.2	380	43	AI165643	AI165643 A087P60u
13	220.2	11.1	412	79	AM310334	AM310334 sf34f03.x
14	198.4	10.9	474	46	AI416897	AI416897 sal9e11.x
15	186.6	10.8	524	46	AM033961	AM033961 EST277623
16	194.8	10.7	475	80	AM329855	AM329855 N201132e
17	181.4	10.0	299	20	D25047	D25047 R1CR2976A R
18	180.2	9.9	714	92	AO969360	AO969360 LERUM53TR
19	178.6	9.8	490	47	AI486325	AI486325 EST244646
20	172.6	9.5	1202	84	B09168	B09168 T2E10-r7 TA
21	159	8.8	1434	63	AI967314	AI967314 U11mpes
22	158.6	8.7	383	81	AM423485	AM423485 sh67e01.Y
23	158.4	8.7	631	61	AM067518	AM067518 660013E08
24	149.8	8.3	647	51	AI730535	AI730535 BNLGH1700
25	145.8	8.0	621	92	AO969361	AO969361 LERUM53TR
26	144.4	8.0	550	74	AM223870	AM223870 EST300681
27	139.8	7.7	548	47	AI728857	AI728857 BNLGH1118
28	139	7.7	608	79	AM234133	AM234133 sf34f03.Y
29	138	7.6	496	62	CNS0079	AL096099 Arabidops
30	134.2	7.4	466	83	AI996242	AI996242 701550775
31	133.2	7.3	385	49	AU056335	AU056335 AU056335
32	133.4	7.3	717	49	AU056601	AU056601 AU056601
33	129.4	7.1	536	62	AI896183	AI896183 EST265626
34	128	7.1	428	28	AM080670	AM080670 EST012 Su
35	125.8	6.9	636	23	T45217	T45217 8480 Lambda
36	125.4	6.9	572	80	AM284352	AM284352 LG1.275.D
37	125	6.9	1345	63	AI967315	AI967315 U11mpes
38	124.8	6.9	550	64	AM042251	AM042251 614026B11
39	124.2	6.8	747	91	AO840411	AO840411 nbx0058F
40	123.2	6.8	477	74	AM224241	AM224241 EST300968
41	122.6	6.8	304	35	C22490	C22490 C22490 Rice
42	122.4	6.7	679	63	AI937984	AI937984 sc06e07.Y
43	120.8	6.7	753	74	AM201125	AM201125 se98B06.Y
44	120.6	6.7	753	105	AO271213	AO271213 nbx00251I
45	119.8	6.6	606	51	AI729170	AI729170 BNLGH1128

## ALIGNMENTS

```

RESULT 1
AA738547 656 bp mRNA EST 17-AUG-1998
LOCUS SBRK5 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
DEFINITION SBRK5 similar to protein kinase, highest similarity to
receptor-like protein kinases, mRNA sequence.
ACCESSION AA738547
VERSION AA738547.1 GI:3421472
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum
REFERENCE 1 (bases 1 to 656)
AUTHORS Annen, F., Chang, J.-D., Paterson, A.H. and Stockhaus, J.
TITLE Characterization of 14 different putative protein kinase cDNA
clones of the C4 plant Sorghum bicolor
JOURNAL Mol. Gen. Genet. 259 (1), 115-122 (1998)
COMMENT 98409267
On Nov 29, 1993 this sequence version replaced gi:636097.

```

Contact: Stockhaus J  
 Institut fuer Entwicklungsbiologie und Molekularbiologie der  
 Pflanzen  
 Heinrich-Heine-Universitaet  
 Universitaetsstrasse 1, 40225 Duesseldorf, Germany  
 Tel: (49)-211-81-14953  
 Fax: (49)-211-81-14871  
 Email: stockha@uni-duesseldorf.de

Insert Length: 1100 Std Error: 0.00.  
 Location/Qualifiers  
 1..656  
 /organism="Sorghum bicolor"  
 /strain="cv. TX430"  
 /db\_xref="taxon:4558"  
 /clone="SBRK5"  
 /clone\_lib="Sorghum bicolor cv. TX430 leaf"  
 /dev\_stage="green"  
 /note="Organ: leaf; Vector: Lambda ZAP II; Site\_1: Eco RI;  
 Site\_2: Xho I; leaf cDNA library from green leaves,  
 unidirectionally cloned"

BASE COUNT 161 a 145 c 184 g 166 t

ORIGIN

Query Match 22.0%; Score 398.4; DB 38; Length 656;  
 Best local Similarity 76.5%; Pred. No. 5.3e-84;  
 Matches 502; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

```

Qy 961 ctacgtggttctcgtacgaccacagagcgcttctgtatcatcatagctaat 1020
Db 1 CTTCGGTTCCTGCTGACGCGCTACTGACCGTGTAGTACCATGAGCTACT 60

Qy 1021 ggaagtggtcgtggttaagaagcgctgacatgaacccctcctgattgg--- 1078
Db 61 GGAGGTGTGACATCACGTTTACGAGAGCGACGACATGACACCTGTTAAGTGGAA 120

Qy 1078 ccaacagaaagagatgctacatagatcgttaaggggcttcttatttcatgacct 1137
Db 121 ACACAGAAAGAGACGATTCGCTGATGATGCAAGAGGACTTCTTACTGTGATGATC 180

Qy 1138 tgtatcccaagatattccatcgtatgataaagctgcaatatatttggagaagaa 1197
Db 181 TGGATCCCAAAATCATTCATCGGGATCTCAAAAGTGCMAAATCTTTTCATAGAGAC 240

Qy 1198 ttgaagcgctgtgaagtatttgggttagctagctgaatgattataagaagataccat 1257
Db 241 TTGAGAGCAAGTGTGGGTATTTTGGGCTTCCCAAGCTTATGACTACAAAGATCCCAT 300

Qy 1258 gttaacatgctgttaaggggtacacttgggtacatagctcccgagttacctcgatgga 1317
Db 301 GTCAACAATGCTGTGCTCGTGAACAAATGACACATTTGCCCTGAGTACTATCCACTGGC 360

Qy 1318 aagtcacagaagaacgcagatgcttgggtatgggattatgctttagagcttact 1377
Db 361 AAGCTCTCGAAGAAAGCATGATTTTGGCTATGGATCATCTCTCGAGTTTATTTACT 420

Qy 1378 ggaacagagcctttagatcttgccttgcgccttgcgaacagatgattatggttggat 1437
Db 421 GGCCAGAGGGCATTTATCTTGGCTGCTCTTGCAGAAATGATGACAGATTTATGCTTTGAC 480

Qy 1438 tgggttaaaagccttttgaagaagaaagttggaagtctgtgcatcctgacctggag 1497
Db 481 TGGGTGAAGGATGCTGTAAGAGGAAGAGAGTGAGATCTGTGGACCAAGATCTGGAG 540

Qy 1498 aacaattacattgacacagaagtgtgagcagcttatcaagtgcattactgtgaccaa 1557
Db 541 AAGCCTCAGAGAGATTCAGAGGTGAGGAACCTGATCCAGCTGCACCTCTGCACGAG 600

Qy 1558 ggttcgccaatgagagcgctaaagtatgataaggttgcgaatccttgaagtga 1613
Db 601 GGCTCCCGCTTGACCTCCCAAGATGTGAGAGGTGGTGGTGGATGCTCGAAGGTA 656

```



```

RESULT 2
LOCUS AA738544 991 bp mRNA EST 17-AUG-1998
DEFINITION SBRLK2 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone
SBRLK2 similar to protein kinase, highest similarity to
receptor-like protein kinases, mRNA sequence.
ACCESSION AA738544
VERSION AA738544.1 GI:3421469
SOURCE EST.
ORGANISM sorghum.
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum.
1 (bases 1 to 991)
REFERENCE 1
AUTHORS Annen, F., Chang, J.-L., Paterson, A.H. and Stockhaus, J.
TITLE Characterization of 14 different putative protein kinase cDNA
JOURNAL clones of the C4 plant Sorghum bicolor
MEDLINE Mol. Gen. Genet. 259 (1), 115-122 (1998)
COMMENT 98409267
On Nov 29, 1993 this sequence version replaced gi:636080.
Contact: Stockhaus J
Institut fuer Entwicklungsbiologie und Molekularbiologie der
Pflanzen
Heinrich-Heine-Universitaet
Universitaetsstrasse 1, 40225 Duesseldorf, Germany
Tel: (49)-211-81-14953
Fax: (49)-211-81-14871.
Email: stockhaus@uni-duesseldorf.de

FEATURES
source
Insert Length: 1500 Std Error: 0.00.
Location/Qualifiers
1..991
/organism="Sorghum bicolor"
/strain="cv. TX430"
/db_xref="taxon:4558"
/clone="SBRLK2"
/clone_lib="Sorghum bicolor cv. TX430 leaf"
/dev_stage="green"
/note="Organ: leaf; Vector: lambda ZAP II; Site.1: Eco RI;
Site.2: Xho I; leaf cDNA library from green leaves,
unidirectionally cloned"

BASE COUNT 253 a 194 c 293 g 251 t
ORIGIN
Query Match 20.1%; Score 365.6; DB 38; Length 991;
Best Local Similarity 64.4%; Fred. No. 3.1e-76;
Matches 599; Conservative 0; Mismatches 319; Indels 12; Gaps 3;

QY 734 ttgttcaactgaagaggtttctctcgtcagaaattgcaagtcgcaacgatactttagta 793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17 TTGGCCAGTTGAAAAGATTTCAGAGAGAGATTGCACACTTCGACATGAAATTTCAGTG 76
QY 794 ccata-----cttggaagaggttgatttgtaagtgataaaggagcgccttgctgaty 847
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 AGAAAATGTTCTTGACAGCAAGGGGCTTGTGGAAAGATATTAAGGACGACTTCACAGATG 136
QY 848 gctcaactgtgacaggttaaaaggcttaaaagaacgaacacgaagttgtagctgcaat 907
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 GCATTAAGATTGCTGTAAACGGTTAACTGATTTTCGAAAGTCTCGGTGGAGAGCTGCTT 196
QY 908 ttcaaacagaggttgaaatgattagcattgctgacatcgaacatctctgctcactgty 967
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 TCTTGCGAGAGTGTAGAGTGATGTGAGAGTTCACCGGAATCTTTAAGATGATG 256
QY 968 gttctgcatgacacacagagagcgtctctgtatatacatcatgagcgaatgagagty 1027
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 GTTCTGTACAAACAAACAGAGCGCTGCTGTGTATCTTTATGACACAACTTAGTG 316
QY 1028 ttgcgtcgtgtttaagagcgtacacatcgaacctcccttgattggtccactagga 1087
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 317 TT---TCTACGCTACAGGAATTTAACCTGGGAGGACAAATATTAGATTGCTGCAAGA 373
QY 1088 agaggaatgcactagatcctcgtcagggtcttcttatttcatacgaatcgtatccca 1147
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 AGCGAGTGGCTTAAGACAGACAGCTCGTGAGACTGAGATTGTCACAGACATGCATCTTA 433
QY 1148 agattatccctgttgtaaaagctgcaaatatattatgtgacgaagaataggty 1207
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 AGATTATACATCGTGAACGTCAAGCTGCCAATGTCTTGCTTATGAAAGGTTTGAACGG 493
QY 1208 ttgtagtgattttgggttagtagcgtcatgattacacagaataccatgtaacatg 1267
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 TTGTGTGATTTTCGGCTTGCCACACTGCTGGATGATACAGAAAGCACTGTGACTACTC 553
QY 1268 ctgtaagggttaccttgggtacataagctcccgagtaacctctcgactggaagtcatc 1327
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 AGGTCCGTGAAACAAATGGGACATTTGCCCTCGAATTTGCCACTGGGAAGTATCCG 613
QY 1328 agaagacagatgctcgttgatgtaggattatgctcttagagctcatatgacagagag 1387
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 AGAGAACCATGTTTGTGTTGCTTACGCAATATGTTCTCGAGCTAGTCACTGTCAGCGAG 673
QY 1388 ctttgatcttgctcgcccttgacgaacatgatatgtatgttgatgattggtttaa 1447
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 CCATGACACTTTTCACGCTGAGAGAGAAAGATGATG---TGTACTTGTATCATGTCAACA 730
QY 1448 gcccttgaaagaaagaatttgagatgctggtgatcctcgaactcgaggaacattca 1507
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 ACGTGCAAAAGGAGGACATCTGACGCCATATGTTACCGTAACGTAACCAACAAATTACA 790
QY 1508 ttgacacagaagttagagcaagcttatcaagtagcaattactctgtacccaggttgccaa 1567
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 ACGGCGAGAGAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850
QY 1568 ttgagcggccttaagatgacagagtagtgcgaatgcttgaagtgatgagccttgcaaa 1627
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 AGGACCGCGCGTCCATGCTCCAGAGTGGTCCGATGCTGGAAGGCGAGGCTTCGACAGAG 910
QY 1628 agtggacagagtgacaaagaatgaagttca 1657
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 GGTGGAGAGAGTGGACAGCAGCTGAGAGTGA 940

RESULT 3
LOCUS AM233982 532 bp mRNA EST 13-DEC-1999
DEFINITION sf32g05.v1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl02e-17123 5' similar to TR:023921 023921 SOMATIC EMBRYOGENESIS
RECEPTOR-LIKE KINASE.; mRNA sequence.
ACCESSION AM233982
VERSION AM233982.1 GI:6566309
SOURCE EST.
ORGANISM soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Glycine.
1 (bases 1 to 532)
REFERENCE 1
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gleibson, M., Pape, D., Harvey, N.,
Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterson, R. and Wilson, R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT On Feb 18, 1993 this sequence version replaced gi:4297707.
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```



Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134 For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or  
 info@genomesystems.com web site: www.genomesystems.com  
 High quality sequence stop: 403.

## FEATURES

## SOURCE

1..532  
 Location/Qualifiers  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1713"  
 /clone\_lib="Gm-cl028"  
 /tissue\_type="roots of 'superpod' plants"  
 /lab\_host="DH10B"  
 /note="Vector: Bluescript II XR, Site-1: EcoRI, Site-2:  
 XhoI; The mRNA was isolated from roots of Glycine max  
 'superpod' plants generously donated by Dr. Gary Stacey.  
 The seedlings were inoculated with Bradyrhizobium  
 japonicus, strain USDA110 prior to harvest. Stragene's  
 cDNA synthesis kit (catalog number 200401) was used to  
 synthesize the cDNA. First-strand synthesis was performed  
 with 5-methyl dCTP, hence the ligated cDNA was  
 hemimethylated. A modification of Stragene's  
 first-strand synthesis primer was used. An 'anchor'  
 nucleotide (V-A-C, or G) was added to the 3' end of the  
 primer [GAGAGAGAGAGAGAGAGAGACTACTCTCGAG(T)18V] to anchor  
 the primer at the 5' end of the poly(A) tract. After  
 second-strand synthesis, the cDNA ends were filled in with  
 cloned Pfu DNA polymerase, ligated to EcoRI adapters and  
 subsequently phosphorylated. The XhoI site within the  
 first-strand synthesis primer was then restricted to  
 digest with XhoI: all XhoI sites in the cDNA would be  
 protected by their hemimethylated status. The cDNA  
 constructs were size-fractionated with a 500bp cutoff,  
 using GibcoBRL Life Technologies' cDNA size fractionation  
 column. The column eluent was then ligated into  
 Stragene's Bluescript II XR Predigested vector  
 (Bluescript II SK(+)) that has been digested with EcoRI  
 and XhoI, and phosphorylated by Stragene). Both the  
 white and blue colonies appear to contain recombinant  
 plasmids with cDNA inserts, based on size (n=25). This  
 library was constructed by Dr. Paul Kelm and Dr. Virginia  
 Coreyell."

BASE COUNT 145 a 92 c 129 g 165 t 1 others  
 ORIGIN

## Query Match

Best Local Similarity 80.0%; Score 361; DB 79; Length 532;  
 Matches 424; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 943 catcgaaatctctcgctcgaatggttctgcatacacaacagagcgcctctgtta 1002  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 2 CATCGTAATTCTCGCTCGCTTCGTGTTTGTATGACACCTACGACGGTGTGTG 61  
 QY 1003 tatccatacagtcgaatgaagtgttcgtcgttttaagaagcgcacccatcagaa 1062  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 62 TATCTTACATGCTACGAGAGGTAGCATATCTTTACGAGAACGTCAGAAATCCCA 121  
 QY 1063 cctcccttgatctggccaactgaagagatgtgcaactgaatctctgaagggctcttc 1122  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 122 CCGCCACTTGGCTGGCCAGAACGAGCATTTGCATTGGCATCTGCAAGGGGCTTCT 181  
 QY 1123 tatttgacagaccatttgatcccaagattatccatcgtgatgtaaaagcgcgaatata 1182  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 182 TATTTGCATGATCATTTGAGACCTTAAGATTATTCACCGTGATGTAAAGCAGCTAATATA 241  
 QY 1183 ttattgagaaagattgaagcgtcttgatgtaatttggttgagcgaagtcagat 1242  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 242 TTGTTGAGAGAGAAATTGAACAGATGTTGTGAGATTTGGTTTACCTTAACCTTATGAT 301

QY 1243 tacaagataccatggttacactgctgtaaggggtacattggtggtacatagctcccgag 1302  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 302 TATTAAGATACATCATGTACTACTGCTGTAGTGAAACAATTGGACATATAGCACACGAA 361  
 QY 1303 tacctctgactgtgaagatcacaagaagaccgagtccttggttatggtatgatactg 1362  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 362 TACCTCTCAACTGGAAGACTTCACAGAGACGTGATGTTTGGATATGAGTGTATGCTT 421  
 QY 1363 tttagagctcattactgagcagagagcctttgatctgctgccttgagcgaagatgatag 1422  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 422 CTGGAAGTAAATACCTGGACAAAGGCGCTTTGATCTACTGCACTTGCACATGATGAT 481  
 QY 1423 gttatgcttgatgagtgtaaaagcctttgaaagaaagaaagttgga 1472  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 482 GTCATGTTCTGATTTGGGTTAAAGAGCTTCACAGACAGGAATATGGA 531

## RESULT 4

AI728030 555 bp mRNA EST 11-JUN-1999  
 LOCUS  
 DEFINITION BNGH19629 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar  
 to (U93048) somatic embryogenesis receptor-like kinase [Daucus  
 carota], mRNA sequence.

ACCESSION AI728030  
 VERSION AI728030.1 GI:5046816  
 KEYWORDS EST.

SOURCE  
 ORGANISM Gossypium hirsutum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

## REFERENCE

AUTORS Blewitt,M., Matz,E.C., Davy,D.F. and Burr,B.  
 TITLE ESTs from developing cotton fiber  
 JOURNAL Unpublished (1998)  
 COMMENT On Jun 5, 1998 this sequence version replaced gi:3187310.

## FEATURES

## source

Location/Qualifiers

1..555  
 /organism="Gossypium hirsutum"  
 /cultivar="Acacia Maxxa"  
 /db\_xref="taxon:3635"  
 /clone\_lib="Six-day Cotton fiber"  
 /tissue\_type="Immature fiber"  
 /dev\_stage="Six days post anthesis"  
 /lab\_host="XLI-Blue"  
 /note="Vector: pBluescript II KS+"

BASE COUNT 146 a 100 c 147 g 162 t  
 ORIGIN

## Query Match

Best Local Similarity 77.5%; Score 352.6; DB 51; Length 555;  
 Matches 427; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1205 ctgttgtaagtgatttggttagctagcctcatgattacaagaatataccatgtaaca 1264  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 1 CTGTTGTGTTGACTTTGGGTTGCTAAACTTAATGAGTACAAAGATACCATGTAACATA 60  
 QY 1265 ctctgttaaggggtacactcgtggtctacatagctccgagatcactcgcagatgaagtc 1324  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 61 CTGCTGACGTGGCACAATTTGACATTTGCTCCAGATATCTCTACTGGAATAATCTT 120  
 QY 1325 cagagaagaccgagtgcttggtgatgagattatgctcttagagcctcattactgacaga 1384  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 121 CAGGAAAACTGATGTTTGGGTATGATATCGCTTTTGGAGCTTAATACTGACACG 180



FEATURES	SOURCE
BASE COUNT	365 a 300 c 389 g 381 t
FEATURES	<p>Insert Length: 1900 Std Error: 0.00.</p> <p>Location/Qualifiers</p> <p>1..1435</p> <p>/organism="Sorghum bicolor"</p> <p>/strain="cv. TX430"</p> <p>/db_xref="taxon:4558"</p> <p>/clone="SbRLK3"</p> <p>/clone.lib="Sorghum bicolor cv. TX430 leaf"</p> <p>/dev_stage="green"</p> <p>/note="Organ: leaf; Vector: Lambda ZAP II; Site_1: Eco RI Site_2: Xho I; Leaf cDNA library from green leaves, undirectionally cloned"</p>

ORIGIN

Query Match	19.2%	Score 349	DB 38	Length 1435
Best Local Similarity	64.5%	Prod. NO. 2.6e-72		
Matches 555	Conservative 0	Mismatches 300	Indels 6	Gaps 2
Qy 797	tacttgaaagaggtgatttgtaaggtgataaaggagcgcctgctgtaagtcacttg	856		
Db 530	TTCTTGACACAGGGGCTTTGGGAAAGATATATAAGGACACTCCAGATGCACATAA	589		
Qy 857	tagcaagttaaaggtctaaagagaagaacacgaaggtgtgagagctgcagttcaaacg	916		
Db 590	TTGCTGTAAACGGTAACTGATTCGAAAGTCTCGTGGAGAGGCGCTTCTTTCGCG	649		
Qy 917	aagtggaaatgatgaatgcatgagctgtgacatcgaaactctgctgctcagtggtttctga	976		
Db 650	AGGTTGAGCTGATTAAGTGTGACAGTTCACCGGAACCTTTTAAAGTATATGTTCTGTGA	709		
Qy 977	tgaacaaagaagagcgagctctctgtatataccatagctgaagagagtggtgcgt	1036		
Db 710	CAACACAAAGAGCCCGCTGCTTTGTTTCTTCATGACAAATCTTAAGTGGCTTACC	769		
Qy 1037	gtttaagagcgctcagccatagaacccctccctgattggtgccaatagaagaattg	1096		
Db 770	GTCACACAGA--ATTAACTGGAGGCCAATAATTGATTTGCTGCAAGGAGACCGATGG	826		
Qy 1097	cactagatctgctaaaggggcttctctattgtgacatctgtaacccaagaattacc	1156		
Db 827	CTATAGGCAACAGCTCGTGAGCTGGAGTATTTCACAGACACTGCATCAATGAATTATAC	886		
Qy 1157	atcgtgctgtaaaagctgcgaataatattatggaacgaagaattggagctgttgatg	1216		
Db 867	ATCGTACGTCACAGGCTGCCAAATGCTTGCTTGATGAAGGTTTGAACCGGTTGTTGG	946		
Qy 1217	attctggattagatgaagcctcaatgattcaagaataccatgctacaacatgctgaagg	1276		
Db 947	ATTTCGGCTTGGCCCAAGCTGTGATGATGACAGACATCTGTGCTGCTCAGGTCCG	1006		
Qy 1277	gtacctggagctcatagctcccgagtaacctctgcagctggaaatctcaagagaagccg	1336		
Db 1007	GAACCTATGGTCACTATGGCCCCGATATATTTCTCCACGTGGGAATGCAATCCGAGAAACCG	1066		
Qy 1337	atgctcttggttattgggaattatgctctcttagagctcatattaccggaacagaagctttgata	1396		
Db 1067	ATGTTTGTGGTATCGGCATAATACCTTCTGACACGTGTGACGTGCACGTCGATAGAC	1126		
Qy 1397	tgtgcgccttggaagacggttgatgattgattgtgttgattggttgtaaaagcctttga	1456		
Db 1127	TTTCACTCTGGAGAGAGAGATGATG--TGTACTTGATCATGTCAACACAGTCGAA	1183		
Qy 1457	aagaagaaaagttgagatgctggtgcatactgacactggaagaaatatacttgaacag	1516		
Db 1184	GGGAGAGGCACTCCGACGCCATAGTGAACCCGTAACCTGAACAAATAATTAACACGGGAGG	1243		
Qy 1517	aagtgtgagcagctattatcaagtagcacttctctgtgatacccaaggtttgcgaatgagagcg	1576		
Db 1244	AGGTGGAGATGATGATCCAGATCGGCTCTCTGCAACCGACAGCTTCGCTTGAGGACCGCC	1303		
Qy 1577	ctaagatgtcagaaggtagctccgaatgcttgaaggtgatgagtccttgacgaaaagtggagcg	1636		
Db 1304	CGTCCATGCTCCGAGGTGCTGCCGATCTTGGAAGGCGAGGCGCTCGCAGAGAGTGGGAGG	1363		
Qy 1637	aagtgcataaagttagaatca 1657			
Db 1364	AGTGCAAGCAAGTGGAGGTGA 1384			
RESULT 6				
AI900110	538 bp	mrna	EST	06-DEC-1999
LOCUS				
DEFINITION	sc01042.y1.Gm.c1012	glycine max cDNA clone	GENOME SYSTEMS	ID
	Gm.c1012-871.5'	similar to TR:023921	SOMATIC EMBRYOGENESIS	



RECEPTOR-LIKE KINASE. ; mRNA sequence.  
 ACCESSION AI900110  
 VERSION AI900110.1 GI:5606012  
 EST.  
 KEYWORDS soybean.  
 SOURCE soybean.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycyne.  
 REFERENCE 1 (bases 1 to 538)  
 AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steplear, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT On Jun 5, 1998 this sequence version replaced gi:3189472.  
 Contact: Shoemaker R./Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com  
 Seq primer: -40RP from Gluco  
 High quality sequence stop: 411.  
 Location/Qualifiers  
 1. 538  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: GM-cl012-871"  
 /clone\_1lb="GM-cl012"  
 /tissue\_type="Apical shoot tips, 9-10 day old etiolated seedlings"  
 /lab\_host="XLI0-Gold"  
 /note="Vector: pBluescript II XR; Site: 1: EcoRI; Site: 2: XhoI; This cDNA library was constructed from mRNA isolated from the apical shoots of 9 to 10 day old etiolated seedlings. The shoot tips including any emerged leaves were harvested for mRNA isolation. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI0-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."  
 BASE COUNT 153 a 105 c 142 g 138 t  
 ORIGIN  
 Query Match 18.3%; Score 332; DB 62; Length 538;  
 Best Local Similarity 77.7%; Pred. No. 2.6e-68;  
 Matches 401; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
 QY 1242 ttacaagatacccatgttacaactgctgaagggtacctgtacatacagtcacca 1301  
 ||||||| ||||| ||||||| ||||| ||||| ||||||| ||||| ||||| |||||  
 DB 1 TTACAAGACACGCATGTGACAACTGCTACGGGGCATATAGCTCTCTGA 60  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1302 gtacttcgactggaagtcatacagagaacgcagtgcttgcttaagcagatagct 1361  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 61 GTACCTTCCACTGGTAAATCTTCAGAGAAACTGACGTTTGGTTATGATATCATGCT 120  
 QY 1362 cttagagctcattactgagacagagagcttttgatctgtcgccttgcygaacgatatga 1421  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 121 TCTGAGACTGATCACTGACAGAGAGAGCTTTTGACCTGCTGCACTGCTAAGACGATGA 180  
 QY 1422 tgtatgtcttgatgagtggttaaaagccctttgaaagagaagaagtgtgagatgctgt 1481  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 181 TGTATGCTCTGATGATGGTTAAAGAGACTTCGAAAGAGAAAAAAGCTGAAATGTTGGT 240  
 QY 1482 cgaaccctgagagacattacattgacacagaagctgagcagcttatacaagtagc 1541  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 241 AGATCTGATCTACAAACCACTATATAGAACTAGAGTAGAACAATATACAGGTTGC 300  
 QY 1542 attactgttaccagaggttcgcgaatgagagcgacctgaagtgcagagtgatgcgaat 1601  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 301 ACATCTGACACACAAAGATTCGCCGATGAGCCACCACTAGATATCAGAAAGTGTGAAT 360  
 QY 1602 gcttgagtgatgagcttccttcagaaagtgagcagctgcaaaagttagatcaca 1661  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 361 GCTTGAAGGTGATGCTTGCGAGAAAGATGGATGAGTGCAAAAGTGGAATTTCCG 420  
 QY 1662 tcaagacgtagaattagctccacatcgaaacttgatgagatccctagactgcagagata 1721  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 421 GCAGAGGTGAGCTGGCCCTCATCTTAATTCGATGATGATGCTGACACCGAAAA 480  
 QY 1722 ctgcagctcttgatattatcttgatcagagataac 1757  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 481 TCTGATGCACTAGATGATTAATCTGCTGCAAGGTGACC 516  
 RESULT 7  
 AI896277 555 bp mRNA EST 27-JUL-1999  
 LOCUS EST265720 tomato callus, TMU Lycopersicon esculentum cDNA clone  
 DEFINITION cLECI4J5, mRNA sequence.  
 ACCESSION AI896277  
 VERSION AI896277.1 GI:5602179  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 555)  
 REFERENCE Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, L.E., Liang, F., Opton, U., Craven, M.B., Bowman, C.L., Ahn, S., Romling, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.  
 TITLE Generation of ESTs from tomato callus tissue  
 JOURNAL Unpublished (1999)  
 COMMENT On May 18, 1998 this sequence version replaced gi:3137451.  
 CONTACT: David Frisch  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 4366  
 Fax: 864 656 4293  
 Email: dfrisch@CLEMSON.EDU  
 5 prime sequence.  
 FEATURES  
 source  
 1. 555  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLECI4J5"  
 /clone\_1lb="tomato callus, TMU"  
 /tissue\_type="callus"  
 /dev\_stage="25-40 days old"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: pBluescript SK(-); Site: 1: EcoRI; Site: 2: XhoI; supplier: Giovannoni laboratory; cLEC - Coryledons







```

VERSION AA738546.1 GI:4242614
KEYWORDS EST.
SOURCE Sorghum.
ORGANISM Sorghum bicolor
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
          Poaceae; Sorghum.
REFERENCE 1 (bases 1 to 655)
AUTHORS Annen, F., Chang, J.-L., Paterson, A.H. and Stockhaus, J.
TITLE Characterization of 14 different putative protein kinase cDNA
          clones of the C4 plant Sorghum bicolor
JOURNAL Mol. Gen. Genet. 259 (1), 115-122 (1998)
MEDLINE 98409267
COMMENT On Jan 13, 1998 this sequence version replaced gi:3421471.
          Contact: Stockhaus J
          Institut fuer Entwicklungsbiologie und Molekularbiologie der
          Pflanzen
          Heinrich-Heine-Universitaet
          Universitaetsstrasse 1, 40225 Duesseldorf, Germany
          Tel: (49)-211-81-14953
          Fax: (49)-211-81-14871
          Email: stockhaus@uni-duesseldorf.de

FEATURES
    source             Insert Length: 600 Std Error: 0.00.
                        Location/Qualifiers
                        1..655
                        /organism="Sorghum bicolor"
                        /strain="cv. TX430"
                        /db_xref="taxon:4558"
                        /clone="SDRLK4"
                        /clone_id="Sorghum bicolor cv. TX430 leaf"
                        /dev_stage="green"
                        /note="Organ: leaf; Vector: Lambda ZAP II; Site 1: Eco RI;
                        Site 2: Xho I; leaf cDNA library from green leaves,
                        unidirectionally cloned"

BASE COUNT    169 a      124 c      179 g      183 t
ORIGIN
Query Match      14.8%; Score 268.2; DB 46; Length 655;
Best Local Similarity 65.3%; Pred. No. 3.1e-53;
Matches 412; Conservative 0; Mismatches 213; Indels 6; Gaps 1;

QY 744 gaagaggttctctcgcgaagattcgaagtcgaagctatttga-----accat 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 GAAAGATTTCGATGCGAGATTCGAACTTCGAACTGATATATTTCAGTGAGAAAAATGT 62

QY 798 acttggaagaggttgattgtgaagggtatagaaggagccttgctgattgctcaattgt 857
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 TCTTGACAAAGGCGCTTTGGGAAAGTATATAAGGAGCACTTCACAGTGCACATAAGAT 122

QY 858 agcagttiaaagggttaagaagaacgaacacccaggtgtgtagctgcaggttcaacaga 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 TGCTTTAAACGCTTAACCTGATTAACGAAGTCTGCTGAGAGGCTGCTTCTTGCGTGA 182

QY 918 ggttggaatgattagcagctgtgcatcgaacatctctcgtcgtcagctgttcgtcat 977
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 133 GGTGAGCTGATTAGTGTTCGAGTTCCACGGGATCTTTAGATTGATGGTTTCTGAC 242

QY 978 gacaccaacagagggctctctgtatcatcacatcagtcataatggaagtgttcgtcgtg 1037
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 AACACAACAGAGCCCTGCTTGTATTATCTTTCATGACGATCTTAAGTGTGGCCTACCG 302

QY 1038 tttaagagagcgtcagcatcgaacgaaccccttgattggcgaactagggaagagattgc 1097
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 TCTACGAGATTTAAACCTGGGGAGCCAAATTTAGATTGGTCTCGAAGGAAGCAAGTGC 362

QY 1098 actaagatcgtcgaaggaggtcttctattcattcgaatgacattgtgaccgaattatcca 1157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 363 TATAGGCACAGCTCGTGGACTGAGTATATGACGACGACATCTTAAGATTATACA 422

QY 1158 tcgtgatgtaaaagctgcaaatatatttgacgaagaattttaggcgtgttgtagtga 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 423 TCGTGACGTCAAGAGCTGCCAATGCTCTGCTTCATGAAAGTTTGAACCGGTGTGTGTA 482
QY 1218 tttaggttagctagctcatatgattacaagatcccatgttacaattgctgtaagg 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 483 TTTCCGCTTGCCCAACGCTGTGGATGTACAGAAACATCTGTGACTACTCAGGTCGTGG 542

QY 1278 taacttggtcacaatagctcccgagctacctcgcagcttgaaagatcataagaagaccga 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 543 AACTAGGCTGCACATTCGCCCTCGAATATTTCACACTGGAAGATCATCCGAGAACCGA 602

QY 1338 tgccttggttatgggataatgccttagag 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 603 TGTTTTGGTTACGGCATATGCTCTCGAG 633

RESULT 10
A1729440 703 bp mRNA EST 11-JUN-1999
LOCUS BNLGH113383 six-day Cotton fiber Gossypium hirsutum cDNA 5' similar
DEFINITION to (U93048) somatic embryogenesis receptor-like kinase [Daucus
          carota], mRNA sequence.
ACCESSION A1729440
VERSION A1729440.1 GI:5048292
KEYWORDS EST.
SOURCE upland cotton.
ORGANISM Gossypium hirsutum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
          eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 703)
AUTHORS Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.
TITLE ESTs from developing cotton fiber
JOURNAL Unpublished (1999)
COMMENT On Jun 5, 1998 this sequence version replaced gi:3188791.
          Contact: Ben Burr
          Biology Department
          Brookhaven National Laboratory
          Upton, NY 11973, USA
          Tel: 516-344-3396
          Fax: 516-344-3407
          Email: burr@bnl.1.bnl.gov
          Seq primer: T3 Primer.

FEATURES
    source             Location/Qualifiers
    1..703
    /organism="Gossypium hirsutum"
    /cultivar="Acacia Maxxa"
    /db_xref="taxon:3635"
    /clone_id="Six-day Cotton fiber"
    /tissue_type="Immature fiber"
    /dev_stage="Six days post anthesis"
    /lab_host="XIL-Blue"
    /note="Vector: pBluescript II KS+"

BASE COUNT    188 a      148 c      180 g      186 t      1 others
ORIGIN
Query Match      13.8%; Score 250.4; DB 51; Length 703;
Best Local Similarity 65.1%; Pred. No. 5e-49;
Matches 368; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1070 ttgattggccaactagaagagatgacatagatcgtcagggaggtcttctatttgc 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9 TACATTGGCCTCAAAAGAAAAGATTAACCTTAGTGTCAGCAGTGGCTGGAGTACCTTC 68

QY 1130 atgacatgtgtagtcccaagattatccatcgtgtaagtaaaagctgcaaatatattgg 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 ATGAACTTGCAATCTTAAGATCATTCACGAGATGTGAAGCTGCAAAATGATTAATGG 128

QY 1190 acgaagaattgagaggtgtgtgtagtgatttgggttgctagctatgattacaagg 1249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 ATGAAGATTTTGAAGCAAGTGTGGTGACTTGGCTTCGGAAGTTGGATGTGAAGC 188

QY 1250 ataccatgttaacaactgtgtaagggtgacttgggtcacaatagctcccgatcctt 1309

```



```

Db 189 GGAAGATGTGACAACTCAAGTTGCGGACAAATGGACACATGACACCGAGTACTGT 248
      || ||| ||||| || ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1310 cgaatgaaagtcctcagagaagccgcatgtcttggttgatattctcttagagc 1369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 CCACTGGGAAGTCTCATGAGAAAGAAAGAGATGCTTTGGTTTATGGATTATCTTCTAGC 308
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1370 tcatctactgacagagagctttgatctgtcgccttcgcaacgatgatgtatgt 1429
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 TTGAGACAGGCAACGTCATTTATTTTCACCCCTTGGAAGATGAGATGATCTTCTGC 368
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1430 tcttgatgggttaaaagccttttgaaagaaaagtggagatgctggtgcattc 1489
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 TTTCTGACTATGTCANNAATCTGAAAGGAGAAACCTGATGCTATGATGATCATA 428
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1490 accctggagaaatattacattgacacagagttgagcagctattccaagtgcattctc 1549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 429 ATCTTAATTAATAATTACACATGAGAGAGCTTAAGCAATGATCAATGATGATGCTTT 488
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1550 gtaccaggttcgcgaatgagcgccttaagatgctcagagtagtccgaatgctgaag 1609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 GCACCCGCGCTCCGCCAAAGAACCTCGGGAATTTTGAAGTGTATGATCTGAGAG 548
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1610 gtgatggccttcgcagaaagtggga 1634
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 GAACAGGCGCTCCGCCAAAGGTGGA 573
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
A1486453 471 bp mRNA EST 29-JUN-1999
LOCUS EST244774 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
DEFINITION CLEB82, mRNA sequence.
ACCESSION A1486453
VERSION A1486453.1 GI:4381824
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 471)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.J., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Roming,C.M., Craven,M.B.,
Fuji,C.Y., Bowman,C.L., Niemman,W., Fraser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
On May 7, 1998 this sequence version replaced gi:3121325.
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.

FEATURES
SOURCE location/Qualifiers
1..471
/organism="Lycopersicon esculentum"
/cultivar="TAA96"
/db_xref="taxon:4081"
/clone="CLEB82"
/tissue-type="carpel"
/dev-stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLEB - Tomato Carpel EST library. Oligo(dT)-primed and
directionally cloned cDNA in vector lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."

BASE COUNT 136 a 88 c 112 g 135 t

```

```

ORIGIN
Query Match 12.3%; Score 223.2; DB 47; Length 471;
Best Local Similarity 67.8%; Pred. No. 1.3e-42;
Matches 312; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy 992 ggcctcttatacatcacaatgagtaagagtgctgcgtggttaagagagcgc 1051
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GCCTGTACTATATACCAATATGCAAGATTTAAGTGCTTATGCTATGCTAGTAACCTA 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1052 agccatcagaacctccctcttgattggccaactagaagaagatgacatgactgcta 1111
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AACCTGGGAGAGCGCTGTTTGAATTGGCCAACTAGGAAACGCTGTGGGCTGACTGCAC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1112 ggggggttctctatttgatgacatgacatgctgacccaagaatatcatgtgataaag 1171
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GTGGACTTAATTAACCTACATGACACTGTAAATCAAAAGATTTATCCGCTGATGTTAAG 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1172 ctgcaatatattattgacagaagaattgagcgtgttgatgatttggttagcta 1231
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CACTAATGTGTACTAGATGAGATGAGATTTTGAAGCTGTAGTGTGATTTGGCTGGCAA 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1232 ggcctatgattacaagaagataccatgattacaactgctgtaagggtaccttggtctaca 1291
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGCTAGTTGACGTTAAGAAACCAATGTGACTACATCAAGTTGCTGTAATAATGGCCATA 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1292 tagctccgaagtaacctctcgtacatgagaagtcacagagaagcagatgcttgattg 1351
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TAGCTCTGAATATCTTATCCACATGGCAATATCAAGAAATCAATGATTTTGGCTATAG 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1352 ggaattatgctcttagagccatcattactgacagagaagctttgatctcgcctgcga 1411
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GATCATGCTTTTGGAAATGTTGAACCGGCCAACGTCATATACCTTCACGCGCTTAGAAG 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1412 acgatgatgatgttattgttctggttgatgggttaaaagcct 1451
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ATGAGATGATGATGCTTGTGCTGACCATGATCAGAAACT 460
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
A1165643 380 bp mRNA EST 03-DEC-1998
LOCUS A087P60u Hybrid aspen plasmid library Populus tremula x Populus
DEFINITION tremuloides cDNA 5', mRNA sequence.
ACCESSION A1165643
VERSION A1165643.1 GI:3856928
KEYWORDS EST.
ORGANISM Populus tremula x Populus tremuloides.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 380)
AUTHORS Sterky,F., Regan,S., Karlsson,J., Hertzberg,M., Rohde,A.,
Holmberg,A., Amin,B., Bhalerao,R., Larsson,M., Villalobos,R., Van
Montagu,M., Sandberg,G., Olsson,O., Teeri,T.T., Boerjan,W.,
Gustafsson,P., Uhlen,M., Sundberg,B. and Lundberg,J.
Gene discovery in the wood-forming tissues of poplar: Analysis of
5,692 expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)
99007314
On Jan 17, 1998 this sequence version replaced gi:2044473.
COMMENT Contact: Sterky F
Department of Biotechnology
Royal Institute of Technology (KTH)
Teknikringen 34, S-100 44 STOCKHOLM, Sweden
Tel: +46 8 790 8287
Fax: +46 8 24 54 52
Email: fredrik@biochem.kth.se
PCR primers
FORWARD: AAAGGGGATGTGCTGCAAGCGG
BACKWARD: GCTTCGGCTGTATGTTGTGTG

```



Seq primer: CGTTGTAACGACGGCCAG  
High quality sequence stop: 380.  
Location/Qualifiers

FEATURES  
source  
1.380  
/organism="Populus tremula x Populus tremuloides"  
/db\_xref="taxon:47664"  
/clone\_lib="Hybrid aspen plasmid library"  
/tissue\_type="Cambial region"  
/dev\_stage="1.5 m actively growing tree"  
/lab\_host="E.coli"  
/note="Vector: Bluescript SK; Site\_1: SalI; Site\_2: NotI; Cambial region tissues, including developing xylem, the meristematic cambial zone and the developing and mature phloem, was harvested from 1.5 m actively growing trees. cDNA was prepared and cloned into lambda gt2a. DNA was isolated and subcloned into Bluescript SK using SalI and NotI restriction enzymes."

BASE COUNT 109 a 67 c 99 g 102 t 3 others  
ORIGIN

Query Match 12.28; Score 220.6; DB 43; Length 380;  
Best Local Similarity 76.7%; Pred. NO. 5.3e-42;  
Matches 293; Conservative 0; Mismatches 87; Indels 2; Gaps 2;

Oy 1162 gatgtaaagctgcaatatattatgtgacgaagaattgagctgtgtgaggtat 1221  
Db 1 GATGTAAAGCTGCAATAATTTTGTGGACAGCAATTCGAGTGTGCGGAGATT 60  
Oy 1222 ggttagctagctcattgattacaaagataccatgttcaactgtctgaagggtac 1281  
Db 61 GGGCTGNCANAAATGATGNACTACAGATGCGATGTCCACACTGCTGT-CGTGTA 119  
Oy 1282 ttgggtacacagctccgagctctgctgacgtggaagcaccagagagaccgatgc 1341  
Db 120 ATAGACACATCGCTCCAGAGTACCTACTACGTGTAATCACTGAGAACTGATGT 179  
Oy 1342 ttgtgtatggtatgctctttagagctcattactggaagaagagcttctgtct 1401  
Db 180 TTTGGGATGTAATGATGATGCTCTGAGCTGATTAATCTGACAGAGCGCTTCGAC 239  
Oy 1402 cgctctggaacagatgattgtattgttggattgggttaaaagcctttgaaag 1461  
Db 240 CGGCTGGCAATGATGACGATGCTATCTTATGGGTAAAGCATCTTCAAGGAG 299  
Oy 1462 aaaaagttggagatgctgctgacccctgacgtggaagaacattacattgacagaagtt 1521  
Db 300 AAAAAGCTNAGAAATGCTATGATCTGATCTCT -CAAAACAAATATGTGAACTGAGG 358  
Oy 1522 gaggagcttattcaagtagcat 1543  
Db 359 GAGCAACTAATCCAGGTGCT 380

RESULT 13  
AM310334/c 412 bp mRNA EST 21-JAN-2000  
LOCUS sf34f03.x1 Gm-cl028 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-cl028-1878 3' similar to TR:023921 023921 SOMATIC EMBRYOGENESIS  
RECEPTOR-LIKE KINASE.; mRNA sequence.  
ACCESSION AM310334  
VERSION AM310334.1 GI:6725935  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae;  
Glycine.  
REFERENCE 1 (bases 1 to 412)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,  
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M.,

TITLE  
JOURNAL  
COMMENT  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ratter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M.,  
McCamn,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
On May 18, 1998 this sequence version replaced gi:3138263.  
Other:ESTs: sf34f03.y1  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estevatron.wustl.edu  
This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134 For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or  
info@genomesystems.com web site: www.genomesystems.com  
Trace considered overall poor quality  
High quality sequence stop: 1.

FEATURES  
source

1.412  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1878"  
/tissue\_type="roots of 'Superpod' plants"  
/lab\_host="DH105"  
/note="Vector: Bluescript II XR; Site\_1: EcoRI; Site\_2: XhoI; The mRNA was isolated from roots of Glycine max 'Superpod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus strain USDA110 prior to harvest. StrataGene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A-C or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGAGACTGTCGAG(18V)] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 107 a 117 c 62 g 125 t 1 others  
ORIGIN

Query Match 11.14; Score 201.2; DB 79; Length 412;  
Best Local Similarity 68.3%; Pred. No. 2.1e-37;  
Matches 278; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Oy 1240 gattacaaggtaccatgttaacagctgtaagggtaccttggtacatagctccc 1299  
Db 412 GATCACACACTACTCATGTTACTACTGTATNCCGTGCGATGCGCAAAATGACAC 353  
Oy 1300 gattacactcgactggaaagctcagagaagaacagatcttggtagtattgatt 1359  
Db 352 GAGTTTCTCACACTGAAAGGCTTTCAGAGAACATGATGTTGGCATATGCAATG 293



at@watson.wustl.edu  
 This library is available through: Genome Systems, Inc. 4633 World  
 Marquette St. Louis, Missouri 63134 For further information  
 (404) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)  
 or contact: clones@genomesystems.com or  
 genomesystems.com web site: www.genomesystems.com  
 reversed clone: similarly on wrong strand  
 err: T7 ET from Amersham.  
 location/Qualifiers  
 1. -474  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl005-165"  
 /clone\_lib="Gm-cl005"  
 /lab\_host="X10-GC19"  
 /note="vector: pLuscript II XR, Site\_1: EcoRI, Site\_2:  
 XbaI. This library was constructed by Dr. Randy C.  
 Shoemaker and Dr. John Epling, USDA-ARS Agronomy  
 Department, G401 Agronomy Hall, Iowa State University,

Email: dirisch@clemson.edu







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 23:41:42 ; Search time 12463.9 Seconds  
(without alignments)  
-318.517 Million cell updates/sec

Title: US-09-180-798-20  
4081  
Sequence: 1 tctagaacctttgcatcat.....tacgaatcatgctctaga 4081

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : GenBank.\*  
1: gb\_dal.\*  
2: gb\_dal.\*  
3: gb\_dal.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_ov.\*  
7: gb\_ov.\*  
8: gb\_ov.\*  
9: gb\_ov.\*  
10: gb\_ov.\*  
11: gb\_ov.\*  
12: gb\_ov.\*  
13: gb\_ov.\*  
14: gb\_ov.\*  
15: gb\_ov.\*  
16: gb\_ov.\*  
17: gb\_ov.\*  
18: gb\_ov.\*  
19: gb\_ov.\*  
20: gb\_ov.\*  
21: gb\_ov.\*  
22: gb\_ov.\*  
23: gb\_ov.\*  
24: gb\_ov.\*  
25: gb\_ov.\*  
26: gb\_ov.\*  
27: gb\_ov.\*  
28: gb\_ov.\*  
29: gb\_ov.\*  
30: gb\_ov.\*  
31: gb\_ov.\*  
32: gb\_ov.\*  
33: gb\_ov.\*  
34: gb\_ov.\*  
35: gb\_ov.\*  
36: gb\_ov.\*  
37: gb\_ov.\*  
38: gb\_ov.\*  
39: gb\_ov.\*  
40: gb\_ov.\*  
41: gb\_ov.\*  
42: gb\_ov.\*  
43: gb\_ov.\*  
44: gb\_ov.\*

45: gb\_ov.\*  
46: gb\_ov.\*  
47: gb\_ov.\*  
48: gb\_ov.\*  
49: gb\_ov.\*  
50: gb\_ov.\*  
51: gb\_ov.\*  
52: gb\_ov.\*  
53: gb\_ov.\*  
54: gb\_ov.\*  
55: gb\_ov.\*  
56: gb\_ov.\*  
57: gb\_ov.\*  
58: gb\_ov.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4081	100.0	4081	5 A67815	A67815 Sequence 20
2	4039.2	99.0	98471	8 AC012654	AC012654 Arabidops
3	540.8	13.3	2089	5 A67827	A67827 Sequence 32
4	470.2	11.5	88401	8 F23M19	AC007454 Arabidops
5	393.8	9.6	96475	7 A67796	AL035678 Arabidops
6	387.4	9.5	6695	5 A67796	A67796 Sequence 1
7	375.6	9.2	65899	50 AC006436	AC006436 Arabidops
8	289.6	7.1	120787	8 AC002292	AC002292 Genomic s
9	277.8	6.8	86722	7 AB012245	AB012245 Arabidops
10	273.2	6.7	86064	7 AB013395	AB013395 Arabidops
11	270.8	6.6	1755	8 DCU93048	U93048 Dancus caro
12	263.2	6.4	84896	7 AB025639	A67797 Sequence 2
13	252.4	6.2	110211	50 AC005170	AB025639 Arabidops
14	244.6	6.0	89779	7 AB005234	AC005170 Arabidops
15	225.2	5.5	134784	8 F17123	AB005234 Arabidops
16	225.2	5.5	134784	8 F17123	AF1600 Sorghum bic
17	148.8	3.6	2278	8 SBRUK1	Y14600 Sorghum bic
18	122.2	3.0	93353	8 A67884	AL034567 Arabidops
19	122.2	3.0	93353	7 A67884	AL021811 Arabidops
20	112.8	2.8	42356	7 AB018120	AB018120 Arabidops
21	112	2.7	2010	8 AF024650	AF024650 Arabidops
22	112	2.7	5438	7 AT284202	Z84202 A. thaliana
23	112	2.7	59376	50 AC006072	AC006072 Arabidops
24	110.4	2.7	80770	7 AB005240	AB005240 Arabidops
25	108	2.6	3176	5 E12705	E12705 Arabidops
26	108	2.6	3176	8 ATU47029	U47029 Arabidops
27	107.2	2.6	94369	8 AC009755	AC009755 Arabidops
28	107.2	2.6	103960	8 ATAC011664	AC011664 Arabidops
29	106.2	2.6	9295	5 E12706	E12706 Arabidops
30	106.2	2.6	9295	7 A67884	D83257 Arabidops
31	106.2	2.6	77040	50 AC004484	AC004484 Arabidops
32	104.6	2.6	75050	8 AC002334	AC002334 Arabidops
33	103.8	2.5	89338	50 AC006068	AC006068 Arabidops
34	101	2.5	78145	7 AB009053	AB009053 Arabidops
35	100.4	2.5	81736	7 AB019235	AB019235 Arabidops
36	99.2	2.4	14313	7 AP000559	AP000559 Oryza sat
37	99.2	2.4	135634	7 AP000391	AP000391 Oryza sat
38	97.4	2.4	98259	8 AC002392	AC002392 Arabidops
39	97.2	2.4	78378	8 AC007260	AC007260 Arabidops
40	97	2.4	80402	7 AB025636	AB025636 Arabidops
41	97	2.4	113567	8 AF058826	AF058826 Arabidops
42	96.8	2.4	68870	50 AC005312	AC005312 Arabidops
43	96.2	2.4	4604	7 LERPGENE	X95269 L. esculentu
44	95.8	2.3	2808	7 RICPRKI	I27821 Oryza sativ
45	93.6	2.3	89029	50 AC006569	AC006569 Arabidops

## ALIGNMENTS



RESULT 1  
 LOCUS A67815 4081 bp DNA PAT 05-MAY-1999  
 DEFINITION Sequence 20 from Patent WO9743427.  
 ACCESSION A67815  
 VERSION A67815.1 GI:4756638  
 KEYWORDS  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.  
 REFERENCE 1 (bases 1 to 4081)  
 De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.  
 TITLE PRODUCTION OF APOMICTIC SEED  
 JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
 CIBA GEIGY AG (CH)  
 FEATURES  
 source Location/Qualifiers  
 1. 4081  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /clone="ARABIDOPSIS SERK GENE"  
 exon 1280..1367  
 exon 1796..1928  
 exon 2014..2085  
 exon 2203..2346  
 exon 2450..2521  
 exon 2617..2688  
 exon 2772..2884  
 exon 3015..3146  
 exon 3305..3646  
 exon 3760..4081  
 BASE COUNT 1120 a 770 c 785 g 1406 t  
 ORIGIN  
 Query Match 100.0%; Score 4081; DB 5; Length 4081;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 4081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 taacttaacttcaataaatacgggattacgaattacttgcattaaattgaaag 540  
 Db 481 TAACCTTACTTTCATTAATAACGGGATTCGATTCATCTTCACTTAAAAATTTGAAAAAG 540  
 QY 541 gtttttacttaaaagaaagggaccacaagagaggttggaccagaaaggg 600  
 Db 541 GTTTTCTCTTAAGAAAAAGGGACCCACAGAGAGGTTTGACCGAGAAAAAGGG 600  
 QY 601 tgcatagccttaagagcttcaactcttaccacaaacccaagcgatgtaacttcaa 660  
 Db 601 TGCATAGCCTTAAGAGCTTTCATCACTTATACCCAAACCAAGCGATGCTACTTCAA 660  
 QY 661 ccactctctctcccccgaaccggttttttggacggttcggtcggaacgaacgct 720  
 Db 661 CCATCTCTCTCTCCCCAACCCGCTTTTGTGACCGGTTCAGTTGCGCACACACCGT 720  
 QY 721 taacggaactataattccctcctcctcctcctcctcctcctcctcctcctcctcct 780  
 Db 721 TAGGGCAGCTTATATCT 780  
 QY 781 cccgttgagatctttaaataataataataataataataataataataataataata 840  
 Db 781 CCGGTGAGATCTTAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 840  
 QY 841 agagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 900  
 Db 841 AGAAGAAAAAGAGAGAC 900  
 QY 901 gtgcggttaataaaagctcttctgtatattacttaccagtagatttcccaaaaa 960  
 Db 901 GTTGGCGTTAAATAAAAGCTCTTTGTATATATATATATATATATATATATATATAT 960  
 QY 961 gctctatttttggtttaaaaaaaagtctcattcatttcaacttttttaagt 1020  
 Db 961 GCTCTATTTTGTGTTAAAAAAAAGTTTCATCTTATTCACCTTTTGTATACGT 1020  
 QY 1021 gtgtgtgaagaagaaggt 1080  
 Db 1021 GTGTGTGTGAAG 1080  
 QY 1081 aattagattttattttatttttttactcttctgttctgtttaaagctaaaggttttaa 1140  
 Db 1081 AATTAGATTTTTATTTTATTTTATCTTTGTTGTTTAAATGCTAATGGGTTTTTAA 1140  
 QY 1141 aagggttaccgaagaagaatgagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1200  
 Db 1141 AAGGTTATCGAAAAAATAGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200  
 QY 1201 tggtagatttcggaagtttaggttttctcgatctgaagagatacaatcaaatcgaaa 1260  
 Db 1201 TGGTAGATTTCGGAAAGTTAGGGTTTCTCGGATCGAAGAGATCAAAATCAAGATTGAAA 1260  
 QY 1261 tttagcatgttgttgaatagagtgagttatgtgtgttttacttacttcaactgat 1320  
 Db 1261 TTTAGCATGTGTTGAAATGAGTGAAGTCAGTTATGTGTGTATCTTACCTTCACTGAT 1320  
 QY 1321 cttaacttcgaatcattcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1380  
 Db 1321 CTTACTTCGAATCATTCACTGTGGGTTGCTCTGCTATTTGGAAGTTGTGTGTGTACT 1380  
 QY 1381 caattactcaagcttactcgttttctcaattacttctcgaattcttttatttggaggtg 1440  
 Db 1381 CAATTACTCAAGCTTCTGTTCTGTCATTAATTCTCTGATTCCTTTTATTTGAGAGTG 1440  
 QY 1441 aatcgcatctttaggt 1500  
 Db 1441 AATCGCATCTTAGGT 1500  
 QY 1501 gtaagatttagtggctagtaacttgaatacaactgtttgtcttttctgttgcagaact 1560  
 Db 1501 GTAAGATTAGTGGCTAGTACTTGAATACACTGTTTGTCTTCTGTTCAGATCAACT 1560



QY 1561 ttgatattgtaaaagcagcttcttggtgtgtaaaagctgggtatttgatattgaa 1620  
|||||  
Db 1561 TTGATATTGTAAGAGCATGTTCTTGGGTGAAGAGCTGGGTATTGATATTGAA 1620  
1621 ttgatgtgtgatcccaaatctcttgaaagactcttggcttgggtttgaa 1660  
|||||  
Db 1621 TTGAGTGTGTGATCCAAACATCTCTGAAAGACTCTATTGGTTTGGTTTAAAGA 1660  
1681 attgtttaattatagctcttaactcagagagagccgtttgaaatgcttcctgaa 1740  
|||||  
Db 1681 ATTGTATTATTTATTTAGCCTTAATCTCAGAGAGCCGTTGAAATAGTCTCTCTGAA 1740  
1741 attagactttcaccaatlgaatgtaattgtgtagattgtgtctgttatagtgat 1800  
|||||  
Db 1741 ATTAGACTTTACCAATTGATGTAAATTGTGTAGATTGTGTCTTGTATAGGTGAT 1800  
1801 gcttgcaactcttgagaggttaactagttgataccaaagtctctgcagagctggat 1860  
|||||  
Db 1801 GCTTGGCACTTTGAGGGTTACTCTAGTTGATCCAAACATGCTTGGCAGAGCTGGGAT 1860  
1861 cctacgctagtgaaatccttgcaactggttcagatgctcaactgcaaaacagagagctgc 1920  
|||||  
Db 1861 CCTAGGCTTAGTGAATCCTTGCAATGTTCCATGCTCACTTGCAACAGAGAAAGTGTG 1920  
1921 ataaagtgtaaagctctctctactaaaccacttttaacttgaactcagcgtggt 1980  
|||||  
Db 1921 ATAAAGGTAAAGCTTCTCTACTAATCCACTTTTAACTTTGACCTGACGCGGT 1980  
1981 taacgaacttttctgtctcttcttgtaaaatagatgattggggaaatgcagagttatctg 2040  
|||||  
Db 1981 TACCAACATTTTGGTTCTTTGTGTCAAATACAGTATTGGGGAATGCAAGATTATCTGG 2040  
2041 ccatttaattccagagctgtgtgtcgaagaatttgcagatttgaatttccactat 2100  
|||||  
Db 2041 CCATTAGTTCAGAGCTGTGTGTCTCAAGATTTGCAAGATTGTGTAAGTTCCACTTAT 2100  
2101 gcatcagcttctaacaacaaatccaaagatttgcagagaagcaactggaattacctt 2160  
|||||  
Db 2101 GCATCATCTTTAAACAAACAAATCCAAAGATTGACAGAAAGACGAGTATACCTTT 2160  
2161 tgaattgaatcttttaacaagttcttatcttaacgggagctttacagtaacaa 2220  
|||||  
Db 2161 TGTAAATGAAATCTTTTAAACAAGTTCTTATTTTCTTACAGGGGCTTTACAGTAACAA 2220  
2221 cataactgcccagatctccagtaactcttgaacttgaacactagtgagttggatct 2280  
|||||  
Db 2221 CATTACTGGCCGATCTCTAGTAATCTTGGAATCTGCAAACTTAAGTAATGGATCT 2280  
2281 ttaactaaacagctctcccggtccactacccgaatcatttgggaaagcttcaaaagctgag 2340  
|||||  
Db 2281 TTACTTAAACAGCTTCTCCGGTCTATTCGGGAATCATTTGGGAAAGCTTTCAAAGCTGAG 2340  
2341 attcttgtagagatacatatgctttacccggtccagttacagttcttggtttaactttagt 2400  
|||||  
Db 2341 ATTTCGTGAGATACATATGTCTTACCGGCTCAAGTACAGCTTGTATATCTTAACT 2400  
2401 ttgttccaacttttgaactcttctgtgtaaaatttacaatgcaagaatgcgggttaca 2460  
|||||  
Db 2401 TTTGTTCCAAATTTTGAAGCTTGTGCTGAATAATTACATGCAAGAAATGCCGGCTTAACA 2460  
2461 acaacagctcagctgggtcaatctcctatgctcagtcagcaaatatactaaccttcaagtgt 2520  
|||||  
Db 2461 ACAAACAGTCTCACTGGGTCAATTTCTATGTCACTACCAATATTTACTCCCTTCAAGGT 2520  
2521 tgtgagctcctcatttaacttcatattgcttacttacttccctcctaagttgattt 2580  
|||||  
Db 2521 TGTGAGTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2580  
2581 gagttaatgcaacttaacttgaatgtaacacacagagagatcatcaaaataacagactctc 2640  
|||||  
Db 2581 GAGTTAATGCACTTAACCTTGAATGATGCAACAGAGATCATTAATAACAGAGCTCTC 2640  
2641 tggttcagttcctcgaaatgctccttctcactcttcaacccatcaggttctatgattt 2700  
|||||

Db 2641 TGGTCAAGTCCCTGCAATAGCGCTCTCACTCTTCCACACCACATCAAGTCTATGATT 2700  
2701 atcctcttaagttattcaagttgtgtgtcagttgtctgaacttattctgaacttcaat 2760  
|||||  
Db 2701 ATCCCTTCAAGTATTTCAGTGTGTGTGCAAGTGTCTCAACTTATTCGAAACCTTTAT 2760  
2761 tcccttgtagctttgtaataacttagacctatgttgacctgtttacaagtcacccatgt 2820  
|||||  
Db 2761 TCCCTTGCAAGTTTGCATTAATTAACCTTAGACCTATGTGACCTGTTACAGTACCCATGT 2820  
2821 ccttgatctcccccgttctctccacacaccttatttcaacctcccccagtttccaac 2880  
|||||  
Db 2821 CCGAGATCTCCCGCTTCTCTCCACACACTTTATTAACCTCCCGAGTTTCCAC 2880  
2881 ccgagaaagcccccctctttaaagttacattatagaaacagagatgaatcttctgt 2940  
|||||  
Db 2881 CCGAGTAAGCTCTCTCTTTTGTGTTACATTAATAGAAACAGAAATGAATTAATCTTGTCT 2940  
2941 cctctgcaatccttctcataactcaactcttgcaataaggaataaccatgatac 3000  
|||||  
Db 2941 CTCTGCAATCCTTTTTCATATAACATCACTTCCATTAAGCAATACCAATATCATC 3000  
3001 taattgattcaagtggtgtaatgataacttggagcaatagctggttggagtgtctgcag 3060  
|||||  
Db 3001 TAATTGATTTGAGGGGATGATGATTAATTAACGACCAATAGCTGATGAGTTGCTGAG 3060  
3061 tgcgtcttgcctcttgcctgcctccgaatgaccttgcgtgtgtggcgaagaagacc 3120  
|||||  
Db 3061 TGGTCTTTGCTCTTGTGCTGCTGCAATTAAGCTTCTTGTGTGCGACAGAAAGCC 3120  
3121 actagatattctctcagatgctcctggtgagtatttacttgcattgattctgttctta 3180  
|||||  
Db 3121 ACTAAGATTTTCTTGATGATGCTGCTGAGTATTAATTTGCAATTAATTTCTGTTCTTA 3180  
3181 gccagcaatttgttctgcagaanaagtalggaaacactgttaatgaaatcaatataca 3240  
|||||  
Db 3181 GCCAGCAATTTTGTGTTGCAAAAGTATGGAACAACTGTAATGAAGAAATCAATACATA 3240  
3241 agtcatgttttttaagtaacaaactctttagaanaactcgtttgcaaatctcat 3300  
|||||  
Db 3241 AGTCATGTTTGTAAAGTATACAAACCTTTTGAATGAATATCGATTCGAAATCTCTAT 3300  
3301 gcaagccgaagaagatccaaagaattcaatctggaacagctccaagaagttcttctgagag 3360  
|||||  
Db 3301 GCAAGCGAAGAGATCCAAAGTTCATCTGGGACAGCTCAAGAGTTCCTTGGGGAG 3360  
3361 ctacaagtggcgaatgagtttagtaacaaagaacatttggcagaagtggtgttgg 3420  
|||||  
Db 3361 CTACAAGTGGCGAGTATGCTTGTAGTAACAAGAACATTTTGGCAGAGTGGGTTGGG 3420  
3421 aaagcttaacaggaagccttggcagaacggaactcttctgtcgttaagaagactgaagaa 3480  
|||||  
Db 3421 AAAGCTTACAGGAGAGCTTGGCAGACGGAACTTGTGCTGTGAAGACTGAAAGAA 3480  
3481 gaagcaactccagtggtgaagctccagtttcaaaagaagtagagatgataagttgca 3540  
|||||  
Db 3481 GAGCGAAGTCCAGGTGGAAGCTCCAGTTTCAAGAGAGTGAATATATATATATGGA 3540  
3541 gttcctogaacactgttgaatgaaggttctglatagacacgacccgaagatgtctt 3600  
|||||  
Db 3541 GTTCATCGAAACCTTGTGAGATTACGAGGTTTCTGTATACACGACGAGAGATGCTT 3600  
3601 gttatccttaacatggcgaatggaatggttggttgtgtctcagaaggttaaaacaaaca 3660  
|||||  
Db 3601 GTGTATCTTACATGCGCAATGGAAGTGTGCTTGTGTGTGAGAGTAAATAACAAACA 3660  
3661 attaaacatcttgcctctctcctaattacttgaagtgatgttttctcaagtttcc 3720  
|||||  
Db 3661 ATTAACATCTTGTGCTCTCTCAATTAATTTGAGGTGAATGTATTTTCAATGTTTCC 3720  
3721 ttataggttcatatatttgtgttacatgaatgaacagagagccacgctcaacactc 3780  
|||||



Db 3721 TTATGGGTGATATTTGGTTACTATGACAGAGAGGCGACCGTACCAACCTC 3780  
QY 3781 cgttgatggccaaacgcggaagaatcgcgcgaagctcagctcagctggttctacc 3840  
Db 3781 CGCTTGATGGCCAAACGGGAGAGAAATCCGGCTAGGCTCGCTCGAGGTTGCTTACC 3840  
QY 3841 tacatgacatgcatgcgaagatcattccacgtgacgttaaaagcagcaaacctctc 3900  
Db 3841 TACATGATCATGCGGATCCGAAGATCATCTCCGTGACGTAAAGCAGCAAACTCTCT 3900  
QY 3901 taagaagaatcgaagcggctgctggaattccgggttgccaagctaatgactata 3960  
Db 3901 TAGACGAAGATTCGAAGCGGTGTGGAGATTCCGGTTGGCAAACTATGACTATA 3960  
QY 3961 aagacactcaatgacaaacagcaatcgtgacaccatcggtcacaatcgtccagaatc 4020  
Db 3961 AAGACACTCAACGTACAAACAGCACTCGGTGACCATCGGCACATGCTCCAAATATC 4020  
QY 4021 tctcaacgcgaatctctcagaagaacacgcgttttcggaatacgaatacgtctag 4080  
Db 4021 TCTCAACCGGAAATCTTCAGAGAAACCGACGTTTCGGATACGGAATCATGCTTAG 4080  
QY 4081 a 4081  
Db 4081 A 4081  
  
RESULT 2  
AC012654 98471 bp DNA PUN 10-DEC-1999  
LOCUS Arabidopsis thaliana chromosome 1 BAC F14023 sequence, complete  
DEFINITION  
AC012654  
VERSION AC012654.2 GI:6554462  
KEYWORDS HTG.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophytes; Magnoliophyta; endicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsi.  
1 (bases 1 to 98471)  
Liu,S.X., Lee,J.M., Sakano,H., Yu,G., Zhaveri,A., Lenz,C.,  
Toriumi,M., Chin,C., Chlou,J., Choi,E., Gonzalez,A., Hwang,B.,  
Koo,T., Li,J., Liu,A., Pham,P., Vaysberg,M., Altafi,H., Brooks,S.,  
Buehler,E., Chao,Q., Conn,L., Conway,A.B., Hansen,N.,  
Johnson,Hopson,C., Khan,S., Kim,C., Lam,B., Nguyen,M., Palm,C.,  
Shum,P., Tambunga,G., Davis,R.W., Ecker,J.R., Federspiel,N.A. and  
Theologis,A.  
TITLE Arabidopsis thaliana chromosome 1 BAC F14023 sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 98471)  
AUTHORS Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (02-NOV-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
REFERENCE 3 (bases 1 to 98471)  
AUTHORS Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (10-DEC-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
COMMENT On Dec 10, 1999 this sequence version replaced g1:615131.  
FEATURES  
location/Qualifiers  
1..98471  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F14023"  
BASE COUNT 31501 a 17757 c 17475 g 31738 t  
ORIGIN

Query Match 99.0%; Score 4039.2; DB 8; Length 98471;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 4070; Conservative 0; Mismatches 3; Indels 9; Gaps 2;  
QY 1 tctagaaacctttgatataatgaataaagagttccatccacacataggggaacat 60  
Db 76556 TCTAGAAACCTTTGATATATGAAATTAAGATGCATCCACACATGGGGTAGAT 76615  
QY 61 aatgtgtatatttaagaggtaacaatgtaactgcttttatttacttttactctc 120  
Db 76616 AATGTGTATATTAAAGGTAAACAAATGTAATCGCTTTTATTTTACTTTACCTCT 76675  
QY 121 actcaaatgtaaggcgag-cttttttttttttaataatgaagaagaatgattta 179  
Db 76676 ACTCAATTTGTTGGCGAGTTTTTTTTTTTTTTAAATGATAAGACAGATGTGTTTA 76735  
QY 180 atgtaattgtaagaagtaglaagtcatalcatalcaggaagcagcatattccaaagt 239  
Db 76736 ATGTAATTGTGATGAACAGTAGTAAGTCAATATGCGGACGCATACTACTCCACAGT 76795  
QY 240 ggaactggcgaatttgccttgcgctcctcacaagtttccaccaatttttgtt 299  
Db 76796 GGAACCTGGCCAAATTTGTCTTCCGCTCTACAGTTCTTCACCAATTTTTTGT 76855  
QY 300 gacaaactcaaatcttccatctcatctctgcgaagttgggttagaagaatcag 359  
Db 76856 GACAAACTCAAACTTTTCAATCTCATCTGCCCCAAAGTTGGGTTTAAAGAAATTCAG 76915  
QY 360 caaacactaatacttatttgttcagttgtatcatcaacaaatcacacaacttgta 419  
Db 76916 CAACACTAATATCTTTATGTTGATGATGATGATCAATCAACAAATTCACAAACATTGTA 76975  
QY 420 aaaaaaatcacaatttctgtaataagatgtaacatgataagaaacttttaacatc 479  
Db 76976 AAAAAAATTCACATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 77035  
QY 480 ttaacttacttcaataatacaggaattacgaacttcttcgtcaattaaatagaanaa 539  
Db 77036 TTAACCTTACTTCTTAATATACGGATTAACGAATCTTTCATTAATAATTTAGAAA 77095  
QY 540 ggttttctacttaagaanaaaggagcccaagagagagtttgacagagaanaagc 599  
Db 77096 GGTTTTCTACTTTAAAGAAAAGGAGCCCAACAGAGAGAGTTTGACGAGAGAAACGG 77155  
QY 600 ggcgaagccttaagaagcttcaacttcaacttcaaccccaacccaagaagatctctca 659  
Db 77156 GTGCATAGGCTTAAAGAGCTTCAACTACTTTAACCCCAAGCCAAAGCATGTCTCACTTCA 77215  
QY 660 accatctctctcccccgaacccgcttttttgacggcgtaagttcgggagagcagcag 719  
Db 77216 ACCATCTCTCTCTCCCGGAACCCGTTTGTGACCGGTCAAGTTCCGGCAGCAGCACCG 77275  
QY 720 ttaaggcagcttataattctcgtctccctccctacacacatgacatgacataataaa 779  
Db 77276 TTAAGGCGAGCTTAATATTCCTGCTCTTCCTCTACACACTGCATCCCATTAATTA 77335  
QY 780 gcccgcttgagatctttaaataataataataataataacgaagaagaatatttattcat 839  
Db 77336 GCCCGTTGAGATCTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 77395  
QY 840 aagaagaagaagaaggaacacacacacacacacatacatagttctctgacagctgt 899  
Db 77396 AAGAAGAAAAGAGAGCAACACACACACACTAATATAGTTTCTGTGGCAGGCTTGT 77455  
QY 900 tctggcgcttaataaagaagctcttcttattattacttcaatgtagattttcccca 959  
Db 77456 TCTTGGCGCTTAATAAAGCTTTGTGATATTAATTAATTAATTAATTAATTAATTAATTA 77515  
QY 960 agctcttatttttggtttaaaaaaaagttctcaacttatttcaactttgttttaag 1019  
Db 77516 AGCTCTATTATTTTGTTTAAAAAAGTTTCAATCTTTATTCAAACTTTTGTGTTTACAG 77575  
QY 1020 tctgtgtgagagagagtggttgattgaggaagcagcagcagagagagcagcagga 1079



D	77576	1GTGTGTGGAAGAGAGAGTGTGATTGAGGAAAGACGACGAGAACCCGGA	77635
O	1080	gaattaggaattcttaatttatttcttactcttctgttcttcaacgtcaatggttttta	1139
D	77636	GAATTAGATTTTGAATTTTATTTTACCTTTGTTTGTTTAAAGCATAAGGGTTTTTA	77695
O	1140	aaagggtatacgaaaaaatgagtgagttgtgttgtagtctgtctctgtcaagttaag	1199
D	77696	AAAGGGTTATCCAAAAAAGAGTGAAGTTGTGTGAGGTTGTCTGTAAATGTTAAAG	77755
O	1200	gtggtgattctcggaagttaggggtttctcgcgatccgaagagatacaaatcaagattcgaa	1259
D	77756	GTGGGATTTTTCGGAAGTAGGGTTTTCGGGATCTGAAGAGATCAAAATCAAGATTTCGA	77815
O	1260	atttagcatctgttcttgaagaatgagtcgagttacgtgtgtttacttacttactcga	1319
D	77816	ATTTAGCATTTGTTGTGAAAGAGAGTCAGTTATGTGTGTTTACTTACTTCACTGA	77875
O	1320	tcttctctccgagatacttaactgtgctgtctctgcctaattggaaggttcgtgttaac	1379
D	77876	TCTTACTTCGAGATCACTTACTGTGGCTTGTCTTCTCTTAATTTGGANAGTGTGTGGTTAC	77935
O	1380	tcaattactcaagcttactacgttcttccaattactctcgaattcttcttactttagaggt	1439
D	77936	TCAATTACTCAGGTTTACTTCGTTTCTCAATTTCTCGATTCTTTTATTTTGGAGGT	77995
O	1440	gaatcgctatcttttagtgcctcgaatttgattatgaaaaattgtgtgttcttttgtatt	1499
D	77996	GAATCCCTATCTTAACTGTGCTGATTTTGATTTATGAAAAATGTTGTGTCTTTGTATAT	78055
O	1500	tctaagaatttaagtgtactacttgaatacaactgttctgtcttctctgtcttcagatcaac	1559
D	78056	TGTAAATTTAGGGCTAGTACTTGAATACACTGTTTTCCTTTCTTGTTCAGATCAAC	78115
O	1560	tctgtatattcgtaaaggcagtctcttgggtgtgaaaagctgggtatttgatactcttaag	1619
D	78116	TTTGTATATTGTAAGGCAATGTTCTTTGGTTGTAAGAGCTGGGTTATTTGATATCTTTAAG	78175
O	1620	attgagtggttgaatcaacaactctctgaagaactcattgttcttggtttgtgaag	1679
D	78176	ATTGATGTTGTTCATCCAAACATTTCTCTGAAGACTTCATTTTGTGTTTGTGTAAG	78235
O	1680	aattgtttaataatataagccctcaatctcaagagaagccgtgttgaaatgctctctctga	1739
D	78236	AATTTGTTTAAATTATTAAGCCTTAATCTCAGAGAGCCGTGTTGAATAGTCTCTCTTGA	78295
O	1740	aattaaacttttaaccaaatgagtcgaattgttgataattgtgtctctgtttaagtgta	1799
D	78296	AAATTAGCTTTTCAACAAATGATGCTAATGTGTAGATTGTGTCTTGTATTAAGTGA	78355
O	1800	tgtcttgcaactcttgaaggttactactagttctgaaccaacaatgctcttgagaagctggga	1859
D	78356	TGCTTTGCACATCTTTGAGGGTTACTGTAGTTGATCCAAACAAATGTCTTGACAGCGGGGA	78415
O	1860	tectaagctagtgaaatcccttgacacatggttccatgtcactcttgacaacagagaacagttct	1919
D	78416	TCTTAGCGTAGTAAATCTTTGCACATGTGTCATGTCACTTGCACAAACAGACAGAGTGT	78475
O	1920	cataaagagtgtaaaagcttctctcaactcccaactttttaaacttgaacctgaagctgga	1979
D	78476	CATTAAGAGTAAAGCTTTTCTTACTACATCCCACTTTTAACTTTGACCTCAGCGTGG	78535
O	1980	ttaaccgacatttctgttctcttctgcaaatacagtgattgtgggaatgacagagttatctg	2039
D	78536	TTTACCGCATTTTGTGTTCTTTTGTCAAAATACAGTGAATTGGGGAAATGCAGAGTATCTG	78595
O	2040	gcaatttaagttccaagagctggtgtgtctcaagaatttgacagattttgtaagttccactta	2099
D	78596	GCCATTTAGTTCCAGACTTGGTGTCTCAAGAAATTTGCAGTATTTTGTAAAGTTCCACTTA	78655
O	2100	tgcatactgtcttaacaaaacaaatccaagaatttgacagagaagaagcacctgagttactct	2159
D	78656	TGCATCAAGCTTTTAAACAAAATAAATCCAAAGTTTGACAGAAAGAACACTGAGTTACTCT	78715
O	2160	tgtgaattgaaatcccttttaacaagttcttatttctcttaacaggagccttaacagtaaca	2219
D	78716	TTTGTAAATGAATCTTTTAAACAAGTTCTCTATTTTCTTACAGGAGACTTTACAGTACAA	78775
O	2220	aataaactgagccagatctcctagtaatactcttgaanaactgcacaacttagtgattgagtc	2279
D	78776	ACATACAGGCCCGGATTCCTAGTATCTTGAAATCTGACAAACTTAGTAGTGGATTGCATC	78835
O	2280	ttaacttaaacagctctccggtccctattccggaatacttgggaagccttcaaaagctga	2339
D	78836	TTTACTTAAACAGCTTCCCGGTCTATTCCGGAATCATTTGGGAAAGCTTTCAAAAGCTGA	78895
O	2340	gatttctgtgagttacataagcttctaccggctgaagttacagttcttgtttaaacttaagg	2399
D	78896	GATTTCTGTGATATACATAAGCTTTTACCGGCTCAGTTACAGTCTTGTGTAAATCTTAAGG	78955
O	2400	tcttgttccaatttcttgactcttctgtctgaaaaatttatactgcaagaataagccggcttaac	2459
D	78956	TTTTGTTCGAATTTTGTGACTGTTTGTGCAAAATTTTACATCGAAGAAATAGCCGGCTTAC	79015
O	2460	aaacaacgttcaactggtgtcaattccctatgtcgaactgaccaatatacctaccctcaagt	2519
D	79016	AAACAACAGTCTCACAGGGGTCAATTCCTATGTCACAGCAAAATATTAACACCTTCAAGTG	79075
O	2520	tgtgtagtccctcaataacttcaatttaattgttcaacttcatctccctcagttgaattgt	2579
D	79076	TTTGTAGTCCTCTCATTTAACTTTATATGTCATCTCATTTCCCTCAGTGAATTTGT	79135
O	2580	tgaatttaatgcaacttaaccttgatgagatgacaacagagatctalcaataacagactct	2639
D	79136	TGAGTTAATAGCACTTAACCTTATGAGAGCAACACAGAGACTTATCAATTAACAGACTCT	79195
O	2640	ctggttcaagttcttaacaagtctctctctcaactcttaaccccaacaggttctataatt	2699
D	79196	CTGGTTCACTTCTTACACATGGCTCTTCTCACTTTCACACCAACAGGTCTATATGAT	79255
O	2700	tatctctctcagttatttcaagtttgtgtgtcagttgtcgaacttattctgaactttcat	2759
D	79256	TATCCCTTCACTTAATTTACAGTTGTTGATGTCAGTGTGCAACTTATTTGAAACTTTTCAT	79315
O	2760	tcccttgtgcagtttctgtcaataaacttagaccatagtgagacctgttaaaagttcacacctg	2819
D	79316	TTTCTTTGTGCAATTTTGTCAATAAATCTTGAACCTATGTGGAACCTGTTTACAGTCAACCCATG	79375
O	2820	tccctgacatcccccgtttctctcccaacaccttttcaaacctcccccagtttcaac	2879
D	79376	TCTGTGATCTCCCCGTTTCTCTCCACACCACTTTTATTCAACTCCCCGAGTTTCCAC	79435
O	2880	cccgaatgaagcctcctcttctttagtttatacatataagaaacagaatgaaatctctgtct	2939
D	79436	CCCGAATAGACCTCCCTTTTATGTTTACATTAATATGAGAAACAGAAATCAAAATCTTTGCT	79495
O	2940	tctctgtcaatcttctttctctataactcactcttggccaataaggaatataaccaatgat	2999
D	79496	TCTCTGTCAATCTTTTCTCTCAATATACATCTTGTCCAAATTAAGCAATTAACCAATGAT	79555



QY 3240 aagtcattgtttttaagttacaacatttttgtagtaaaatctcgatgcaaaactctca 3299  
 |||||  
 Db 79796 AAGTATATGTTTAAATTAAGTTACAACCTTTTGAAGTAAATCTCGATGCAAAATCTCTT 79855  
 |||||  
 QY 3300 tgcagcgaagaagaatccagaagttcatctggagcaagctcaagaagttcttcttgggga 3359  
 |||||  
 Db 79856 TGCACCCAGAGAAAGATCCAGAGATTCACCTGACACTCAAGAGCTTTCTTGGGGG 79915  
 |||||  
 QY 3360 gctcaagttggcgagttgtagtttaagtaacaagaacattttggcgagagttttg 3419  
 |||||  
 Db 79916 GCTAATAAGTGGCGAGTGTATGGCTTATGACAAACAACTTTTGGCGAGAGTGGT 79975  
 |||||  
 QY 3420 gaaagttcaagaaggacgttggcagacgaaactcttggctgctgcaagaagactgaagga 3479  
 |||||  
 Db 79976 GAAAGTCTACAAGGAGGACCTTGGCGAGAGAACTTGTCTGTCAAGAGACTGAAGGA 80035  
 |||||  
 QY 3480 aagaggaactccagttgtagaagctccagtttcaagaagaagtagaagtagaagtaggc 3539  
 |||||  
 Db 80036 AAGAGCAACTCCAGGTGAGAGCTCCAGTTTCACAAACAGAAATGATGATGATGCG 80095  
 |||||  
 QY 3540 agtcatcgaaacactgttggagattcagaagttctgtatgacacgcgacgagagagattgct 3599  
 |||||  
 Db 80096 AGTTTCATGCAAACTGTTGAGATTACGAGGTTCTGTATGACACCGACGAGAGATTGCT 80155  
 |||||  
 QY 3600 tttgtatctcttcaagcgcaatggagagtgcttgcctgtctcagaggttaaaactaac 3659  
 |||||  
 Db 80156 TGTGTATCTCTTACATGCGCAATGGAAGTGTGCTTGTCTGACAGAGTAAAC----- 80211  
 |||||  
 QY 3660 aactaaacacttgcctcctcctcaattacttgaacgtgaagtttttcatgctttc 3719  
 |||||  
 Db 80211 ---TAAACATCTTGTGCTCTCTCTCAATTACTTGAAGTGAAGTGTCTTTCATGTTTC 80267  
 |||||  
 QY 3720 cttatgggttcaataatgttggttaacataatgacagaagggccacggttacaacct 3779  
 |||||  
 Db 80268 CTTTATGGGTTTCAATATTGTTGTTACACTAATGACACAGAGAGCCACCGTCAACACT 80327  
 |||||  
 QY 3780 ccgcttgatggcgcaacgcggaagaatcgcgttaagctcaagctcgaagttgctttac 3839  
 |||||  
 Db 80328 CCGCTTGATGGCGCAACGCGGAAGAGATCCGCGTAGCTCAGCTCGAGGTTTCTCTTAC 80387  
 |||||  
 QY 3840 ctacatgatcactcgatcgatccgaagatcactcaacggtgaacgttaaaagcagaacatccctc 3899  
 |||||  
 Db 80388 CTACATGATCAGTCCGATCCGAGATCTTCACTGACGTAACGTAACGCAACATCTCTC 80447  
 |||||  
 QY 3900 tttagacgaagaattcgaagcgttggtagagaattcgggttggcaagaactaatgactat 3959  
 |||||  
 Db 80448 TTAGACGAAGAATTTCGAACGCGTTTTCGAGATTTCGGGTTTGGCAAGCTAATGACTAT 80507  
 |||||  
 QY 3960 aaagacactcaagtgacaacagcaagtcggtggacacatcggtcaacatcgctcgaatat 4019  
 |||||  
 Db 80508 AAAGACACTCAGGTGACAAACAGACAGTCCGTGACACATCGGTCAACATCGCTCCAGAAATAT 80567  
 |||||  
 QY 4020 ctctcaacggagaatctctcagagaagaacggacgttttcgagatcgagatcatcttcta 4079  
 |||||  
 Db 80568 CTCTCAACGGGAATCTTCAAGAAACCGACGTTTTCGATAGGAAATCATGCTCTTA 80627  
 |||||  
 QY 4080 ga 4081  
 |||||  
 Db 80628 GA 80629

eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.  
 REFERENCE 1 (bases 1 to 2089)  
 AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.  
 TITLE PRODUCTION OF APOMICRITIC SEED  
 JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
 CIBA GEIGY AG (CH)  
 FEATURES  
 SOURCE 1..2089  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /clone="SEER GENE CDNA"  
 195..2072  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAB42254.1"  
 /db\_xref="GI:4756650"  
 /translation="MESSYVVFILSLILPNSHSLMASNLKEDPALHTLRVLVDPN  
 NTVQSDPPLVNPCTWTFHYTCNNNSVIRVDGNALSGHLVPELGVILKNLYELMS  
 NNTGPIPNICNTLNTVSLDXLNSFGSPDESIGKSLRPLRNNSNLGSLPMS  
 LNTITLQVLDLSNNRLSGSVFDNCSFELFTPISRANLIDCGPTSHRCGSPSP  
 PPPPIPPPVSTPSPGIGTGAIVAGVAALPEFAPALFARMMRSPDLDFDVA  
 EEDPEVHLQKRFSLRELVASDEFSNKNILKRGFGKVKYKRLADGTLVAVRLK  
 ERTPGELQFOLEVEIMASVARNILRLGFCMPTEBLVLYPYMANGSVASCLERP  
 PSQPLDMPFKRIALGASGLSYLHDHDPKTIIRDYKAAANILIDEEFEAVVDLGL  
 AKIMDKDHTVTAIVRTGHTLAPELSTGSKSEKTDYGVGIMLELITGQRAFDLA  
 RLANDDVLMDLVKGLERKLEMLVDPDLOTVNEERELDVOVALVITGQSPMER  
 PKMSVYRHLKEDGLAEKRDENQKYLEIRZELDLSPNNSMDILDTYNLNAVELSGP  
 R"  
 BASE COUNT 569 a 419 c 502 g 599 t  
 ORIGIN  
 Query Match 13.3%; Score 540.8; DB 5; Length 2089;  
 Best Local Similarity 85.3%; Pred. No. 1.7e-92;  
 Matches 665; Conservative 0; Mismatches 2; Indels 113; Gaps 1;  
 QY 3302 cagcgaaagatccgaagttcatctggagacagctcaagaagtttcttggcggaagc 3361  
 |||||  
 Db 1018 CAGCGAAGAAAGATCCAGAAATTCATCTGGAGACGTCAGAGGTTTCTTTCGGGAGC 1077  
 |||||  
 QY 3362 tacaagtgcgagtgatgggttttagtaacaagaacatttggcgagaagttgggttggga 3421  
 |||||  
 Db 1078 TACAAGTGGCGAGATGATGGGTTTGTAAACAAGACATTTTGGCGAAGGCGTTTGGGA 1137  
 |||||  
 QY 3422 aagttacaagggacgcttggcagaacgaaactctgttgcctgctcaagagactgaagaag 3481  
 |||||  
 Db 1138 AAGTCTACAAGGAGCGCTTGGCAACGGAACCTGTGTGCTGTCAGAGACTGAAGAAG 1197  
 |||||  
 QY 3482 agcgaaactccagtgtagagctccagtttcaacaagaagttagatgataatgtatggcag 3541  
 |||||  
 Db 1198 ACGGAACCTCCAGGTGAGAGCTCCAGTTTCAMAAAGAAAGTAGATTAAGTATGGCAG 1257  
 |||||  
 QY 3542 ttcatcgaaacactgttggagttacagagtttctgtatgacacgacgacgagagattgcttg 3601  
 |||||  
 Db 1258 TTCAATCGAAACCTGTGAGATTACGAGGTTTCTGTATGACACCGACGAGAGATTCTTG 1317  
 |||||  
 QY 3602 tgtatccttaactggccaatggaagtgtgcttcgtgtctcagaaggttaaaactaaca 3661  
 |||||  
 Db 1318 TGTATCTTACATGCGCAATGGAAGTGTGCTTGTGTCAG----- 1361  
 |||||  
 QY 3662 ttaaacatcttgctctcctcaattacttgcgcggaagtgtttttcatgtttccct 3721  
 |||||  
 Db 1361 ----- 1361  
 QY 3722 ttatggttcataatgttggtttacactaatgacacagagagggccacggttacaacctcc 3781  
 |||||  
 Db 1361 -----AGAGGGCGAACGTCACACACTCC 1384  
 |||||  
 QY 3782 gcttattggccaacgcggaagaagatcgcgctaggtcagctcgaggttggttcttacct 3841  
 |||||  
 Db 1385 GCTTATTGGCCAAAGCGGAAGATCGCGTAGGCTCAGCTCAGAGGTTTGTCTTACTT 1444



QY 3842 acatgctactgctgctccgagatcattcaccgctgacgtctaaagcagcaacatctctt 3901  
|||||  
Db 1445 ACATGATCTGCGGATCCGAGATCTTACCGCTGACGTAAGCAAGCAACATCTCTT 1504  
|||||  
QY 3902 agacagaatctcgaacgctgttggagattcgggttggcaagctaatgactata 3961  
|||||  
Db 1505 AGACGAAGATTCGAGCGGCTGTGGAGATTCGGGTTGGCAAGACTATGACTATA 1564  
|||||  
QY 3962 agacactcagctgaacacagcagctcgtgacacatcgcgtccacatcgtccagaatctt 4021  
|||||  
Db 1565 AGACATCAGCTGACACACACAGCTCGTGACATCGGCATCGCTCCAAATATCT 1624  
|||||  
QY 4022 ctcaaccggaatcttcagagaaacccgcttccgatacgaatcatgctctaga 4081  
|||||  
Db 1625 CTCAACCGGAATCTTCAAGAAACCGACGCTTTCGGATACGGAATCATGCTTCTAGA 1684  
|||||  
RESULT 4  
F23M19/c 88401 bp DNA PLN 17-JUN-1999  
LOCUS Arabidopsis thaliana chromosome 1 BAC F23M19 sequence, complete  
DEFINITION  
ACCESSION AC007454 GI:5091613  
VERSION AC007454  
KEYWORDS HTG.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsis.  
REFERENCE  
AUTHORS 1 (bases 1 to 88401)  
Vysotskaja,V.S., Schwartz,J.R., Yu,G., Toriumi,M., Lenz,C., Liu,S.,  
Lae,J., Liu,A., Li,J., Kremensskaja,I., Luros,J., Gonzalez,A.,  
Altati,H., Ardujo,R., Chao,Q., Conn,L., Conway,A.B., Dunn,P.,  
Hansen,N., Hultzer,L., Kim,C., Palm,C., Kowley,D., Shinn,P.,  
Walker,M., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.  
Arabidopsis thaliana chromosome 1 BAC F23M19 sequence  
Unpublished (1999)  
REFERENCE  
AUTHORS 2 (bases 1 to 88401)  
Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
REFERENCE  
AUTHORS 3 (bases 1 to 88401)  
Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUN-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
REFERENCE  
AUTHORS 4 (bases 1 to 88401)  
Theologis.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUN-1999) Plant Gene Expression Center, 800 Buchanan  
St., Albany, CA 94710, USA  
COMMENT On Jun 17, 1999 this sequence version replaced gi:4996902.  
The sequence of BAC F23M19 from Arabidopsis thaliana chromosome 1.  
FEATURES  
source  
1. 88401  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F23M19"  
/gene="F23M19.1"  
complement(join(796..1116,1205..1271,1349..1432,  
1506..1636,1784..1850,1999..2047,2166..2251,2468..2545,  
2771..2893,2990..3079,3154..3254,3332..3393,3467..3536,  
3638..3891))  
/gene="F23M19.1"  
/note="Contains a pf|00561 alpha/beta hydrolase fold  
domain."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="A039602.1"  
/db\_xref="GI:5091614"  
/translation="MEEDLITSPSPYDLLEFQALFLIPRHRLLAEVLTCTEYNY  
EHEFLDALLFRKRVNLINPDSPLHGVSRKQILHGRVATPLASHPIQICFL  
NFGDLPPVFTYRFLTSDGGTIALDWLNVDVDSGLHNSSELTETPIAVIP  
GLTSDSSAVLKHLADYDARTKGMNVYISNHRGLGVSVYSDCFYAGWTDDIRVLDY  
LOKYPRAPEALGTSIGANVLYKYEAGEKTPRGAVAIICSPMDLIGDFICRTL  
KOKLYDKALITIGIGYALOEHPOLRLAWEGIKRSIRIDGNHATCLVGFKEVDI  
YKRSSTQYGVNAVPIELCISADDPDLCTKRAIIPDEBRANKNIYATVNHGSLAF  
FEGLTSSLMWVAINEFLGVLSCEPTNHIQIVKRSRSGKQEPSINQGYLTAT  
DLVAAVKYEDQNTKNTLKORGRKEEDVTKRSFKELQRTQSVWVLGYIGMTSP  
LGLMLNLYRKORPTTTS"  
complement(5693..6732)  
/gene="F23M19.2"  
complement(join(5693..6108,6233..6428,6499..6732))  
/gene="F23M19.2"  
/note="Similar to gb|X98322 peroxidase (pxr10) from  
Arabidopsis thaliana."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="A039603.1"  
/db\_xref="GI:5091615"  
/translation="KRAITAFEFICFLVPSVFGELRGRGYDSSCPLEASTVRYAK  
HSLNQVYVTRALRMQHDQFVNRKILCVYVGGDASLIDSTPERSEKSAVY  
SVRGFEIIDEKKELEIVCPKTVSCADITVIAKDSIALAGLNTNMVALIGGSHY  
GVAGCSLFQDRILDKPRMDSKRLKQNTORGRPDPSVLDQMTPEVDNQIKQIKSQ  
RGITRIDNGLDDSTSRYSNALNBTLEFGEFPAAMQIMEIVLQNSGEITNC  
IVSH"  
complement(10781..14114)  
/gene="F23M19.3"  
complement(join(10781..11318,11416..11460,11670..11746,  
11875..11991,12207..12310,12443..12512,12614..12717,  
12848..12941,13174..13213,13394..13719,13800..14114))  
/gene="F23M19.3"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F23M19.3"  
/protein\_id="A039604.1"  
/db\_xref="GI:5091616"  
/translation="MGICGRSSVYNNAGCGFAVNHGHINNSDLNHSGEGLK  
DDSPVTENVDNKHTESESPFIYVSGSHPOVTEGICRLSRVLSQKRSKSRQAA  
VAKYSEVSLIGRAGTIGLKAADVLDITGKSNLNLSGGPSATVYGNISLISF  
EVANTIVKGANLHMSLSKDSITLKEVLPSEVQNLISKMDRLRIKAAKRRQD  
PVLIVTQRDLYEHLALDRFEDQYOKIOEENPSTAQGVDTLAIRTELKSGK  
HVRNLKKKSLMSRLTEEMERKLVDVYHFLTEHFAAGDAGDPKAPRPNHKKLS  
AGLAHIANITITDITVSRSSSTMPASTRALYQGIPEPKSLRSRISFOYKEELT  
VPOIKAEKERTLQWLVPAVNTIKAHGRFEGVSWSSGSEANQRPAGQTLKIDTLH  
HADKEKTEAYILDVWLHLVTVQVATVGYGRSPVKSIPRSNPQTIQLSSGSHNP  
SMGLPLTTEDEQMLRDVSKRRTPGISQSEFETVAKRLCHHRLSKSSSHSPMG  
EMKRNKDTFSTRPSPVPIIDPIDRMKALVIDRVDITRL"  
23307..26279  
/gene="F23M19.4"  
join(23307..23355,23431..23627,23656..23751,23850..23906,  
24002..24145,24225..24309,24401..24491,24571..24735,  
24842..24955,25045..25117,25210..25348,25453..25865,  
25953..26140,26231..26279)  
/gene="F23M19.4"  
/note="Similar to gb|AF082176 auxin response factor 9 from  
Arabidopsis thaliana."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="A039615.1"  
/db\_xref="GI:5091627"  
/translation="MESGVNVAQPELGGIIDSRSKYVEQLKLCAGPLCDPKLGE  
KYVYFPGSHIVASLSLSIPFSLSLHSLISLSTVSTREELNLTQPCIDPS  
KLCGRVLAHFKVNNNSDEYATETTPDTQVYIPQNNQRPPLVNSTKYLTLSD  
TSAGGFVPRKTHALECPSLDSKQPLPADELALIDHGQMFNNHNRKTPORHLIT  
TGMNAFTQKSLVAGDIVIVRGTEGLRGIIRRAHQQGNIPSSIVSIDCMHGVYA  
SAKAFDQCMFTVYVPRSKFVYSDKRLDNLNKNFNGFSPRFLRAGDSEBRCK  
FGTIIIGVDFEPHKCSMRSLLEVQNEFISFGPKPKVSPMDIEHLIADNVPSTLL  
KNRLREYNEIGSSSSLLPLPILOGQENQLVASPMNISLYRPAATEDANPSSLL  
MSYVQPMPRKLNQVNTMEENITTKGTNRLRGLVTLDTPTPIKDPTEELGSEIS



KLTEGKFKGLSOTLRSPTEIONKOFSSSRCTKYOKGVYIGRAVDLSYNGDOLIL  
ELEKFDIKGLOLRNOMETIAFTDSDDEKMLVDDDPMPPECNNVKKIFIQKRR"  
complement(28888. .31377)  
/gene="F23M19.5"  
complement(28888. .31377)  
/note="Contains similarity to g1147936 protein kinase PK1  
from Zea mays, is a member of the pf100954 S-locus  
glycoprotein family and contains a pf100069 Eukaryotic  
protein kinase domain."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39605.1"  
/db\_xref="GI:5091617"  
/translation="MAVTPFLKLLPLLLHLHPSEFSTIPLGSYIYASGNOMPS  
PNSTFVSFVSPSPNSFLIAPVAGSVAGSLRLFTSGSLRLTNGSGT  
TYMSKTDRLGVTSGSIDEGEFLILNNRSVPWSSPDNDPTDIYOSOMPTGLIS  
GLYSOLRSGNLTLRNNTSAIYNHGLNLSFSSNLSPLSTOTGVYSTEPSNLG  
CAELIYSGDIYDSDNTFRRLKLDGNTLRISMSRNSGPNAMNSAVDCLYTCGN  
FGICNDNTNDCSPGRNDFVDVNDNRGCKRVELSDSGNTMLDVLHRLFTY  
EDDPNSESFFAGSSPCRANCLSVLCIASVMSDSCNMOKHPSGFTGYOWPVPS  
TSYVVCPPVAANTLERATKGDNNKVLHVIYAVIAGLGLVAVELGLMCCCRK  
NPRGTLSSHYTLLEIYAGAVPTTKEIQRCKRFEKELAGAGFGLVYGLNTY  
YAVKOLEIEEGEKQMEVATISTTHLMVRLIGCSGGRRLIYEPMRNSLDN  
FLTTDSAKFLTWEPYRNIALGTAKGTYLHRCRQCIYHCDIKPENIILVDNPAKY  
SDGLAKLNKNDKRYNMSYRGTGLADPMANLPIRSKYIYCYGATLEIYSGK  
KNDFVSEKTNKRFSLMAYEFENKRALIDRLSLDQSDQVQWRKWTISFWIQ  
KQPLQRPWGVQMLGEITEIKNPLPKTISEFSGNSMSTSHASMPVVASGPTRSS  
SFSATRSFOTGITSVSGPASTRISGSMUGS"  
complement(36027. .36827)  
/gene="F23M19.6"  
complement(36027. .36827)  
/note="Contains two pf100560 leucine-rich repeat domains."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39606.1"  
/db\_xref="GI:5091618"  
/translation="MINRRIYFCLCYWVWVDSRLPYLAIRQVDPIYKIVPIYGR  
FDPEEFTVMQGNNDCEWFGNCLGELIGISFSLNLGTTISHPADLTSLADPS  
HNRLKCTIPEITLTKNLTIYDSYKDLHGVPYRGIVILTRNRISETCLLVDS  
TRNKKRPYVLLGLILGLVYAGASGFGFLIKRQPRLDPEADAVLTLLQOQSSD  
ESLVDSYSVLSLQIKVLRKFSWSKSPILLTRQKTNQNHLPYM"  
40313. .42256  
/gene="F23M19.7"  
join(40313. .40937,41137. .41393,41705. .42256)  
/gene="F23M19.7"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F23M19.7"  
/protein\_id="AAD39607.1"  
/db\_xref="GI:5091619"  
/translation="MVPIASMASTRPLCSIPISFLSFLFVSLLEFPNSGLINP  
NPSISHTLONGINYVAELPSRLNTGIDYWGSSSPDPSDPHPRKHSKD  
YPPYENPLIRQYSAEWIMGDLTSPERKIRGSAKVPSESDAPVPPKPTLSAE  
MEIANGKSPRKSGNEDYQOROVDFYKRTKMKNSGRDHYFYLIDVYAMHRE  
ETALSILYVDFGMEFRQDSKSSNGSLPEHIGTYQSVLKVDFTHLPRLDSQ  
NQRHSLILYFKGAKHRHKGGLIRKLDLWLVNEGVMESGFPAVAREGISIGMNS  
BECLHPADPTSCRLFDALISLCIPVIVSPTLILPEGLIDYSESVFSDALTP  
KMLANHLRFESEREKETLRSHIANYQSVFYVDNGHADGIGPIEPNGAVNHIMKVOOK  
VPMVEAIVIRRRKPAQASVPLRCOCIT"  
complement(43028. .48860)  
/gene="F23M19.8"  
complement(join(43028. .43288,43400. .43672,43778. .45067,  
45136. .45311,45523. .45649,45915. .46249,46361. .47119,  
47284. .47637,47725. .47869,48023. .48613,48703. .48736,  
48815. .48860))  
/gene="F23M19.8"  
/note="Contains similarity to g11836774 FAb1 protein from  
Saccharomyces cerevisiae genome gplD50617."  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="AAD39608.1"

Query Match 11.5%: Score 470.2; DB 8; Length 88401;  
Best Local Similarity 64.5%; Pred. No. 2,1e-79;  
Matches 909; Conservative 0; Mismatches 353; Indels 147; Gaps 8;  
QY 2752 acctcattcttccttgtagcgttttgcctaataactagacctaagtgagacgttgtaacgt 2811  
DB 73879 ACCATTGTTCTTCGACAGATTGCTAACACACTGGATGATGACGCCAGTTACTAGC 73820  
QY 2812 caccatctctgtagctcccccgtttctcccccaccactttttaaccctccca 2871  
DB 73819 GCTCTTGTCTGGAATCCGCCGTTTGTCTCCACACCTTTTAACACCTCCANA 73760  
QY 2872 gtccaccgccagtaaa-----gcctcccttttagt 2904  
DB 73759 GTTCTACACAGGAATGAATGAACAGACAGATAAATTAGACTTTTACTGATTTGACTGCA 73700  
QY 2905 ttacatttaggaacagaagatgaatctttgctctctcgtcaatccctttcattat 2964  
DB 73699 CATGCTTGAAAACATATTGATGTGTATAGAGTTTATCTGATTTGACTGSA 73640  
QY 2965 aactcattctgccaataagccaataacgaatgacttaattg-----gcctcccttttagt 3008  
DB 73639 AAAAGCTATCTACAGATTAAAGACTGTGATCAACTGATGAGTTAATTAATTT 73580  
QY 3008 -----attcaagtggtatggtataactcgtgagcaatagctggtgagtgctc 3054  
DB 73579 TGGTAAATAAATAATTAAAGTGGATGTATGCTACTAGGAGCCATTGCGGAGAGATTGC 73520  
QY 3055 tgcagtgctgctctgctctcttgctgctcctgcatagactctgctggtggtgcgaag 3114  
DB 73519 TCGTGTGCTGCTGTACTATTGCTGCGCCGCTTATGCTTTGCTGTGAGGCTGAAG 73460  
QY 3115 aaagccactagatattctctcgatgctgcctggtgagtttaattatcgaattgctctg 3174  
DB 73459 AAAACCTGAAGAACTTCTTGTGATGTCTGTAAGTCACT-----GACTCG 73411  
QY 3175 tcttcagccaatattgcttttcagaaaagtattggaacaactgtaataaataca 3234  
DB 73410 CAATTCACAGCTTGTCTTCAATTCACAAATTTGAT----- 73374  
QY 3235 tacataagtcattgtttttaaagttacaacactctttgagtaaatcgcattgcaaat 3294  
DB 73374 -----AGATTATGCTACTTAAATTTTTCAGTTCTGT----- 73343  
QY 3295 cctcattagccgaagaaagatccagaagttcaatctggagacgtcaagaggtttcttg 3354  
DB 73343 -----ATGCAGCCGAAGAGACCTGAGGTTCATCTGGGCGACGTTAAGCGGTCTCTCA 73288  
QY 3355 cggagcttacaagtggtgcagtgatggttttagtaacaagaacatttggcagaagtggtg 3414  
DB 73287 CGGGAACCTTCAAGTAGCACTGATGCTTCAGCAACAAGAACATTTTGGGCCGAGGTGG 73228  
QY 3415 ttgggaaagcttaagaagagcgttgtaggaagaaacttttggctgcaagaagcgtg 3474  
DB 73227 TTGCGAAGAGCTTCAAGAGCGCTTCTGTATGGAACACTTGTGAGACCAAGCGCTT 73168  
QY 3475 aaggaagacgaactccagtgtagagactccagtttcaaaacagaagtagatgataagt 3534  
DB 73167 AAAAGAAAGCGAAGCCCGAGTGGGAGGATCCAGTTTCAAGACAGAAGTGAATGATAAG 73108  
QY 3535 atggcaagttaacgaacactgttagatagaggttctgtagaacgacgacgagga 3594  
DB 73107 ATGGCGGTTACAGAAATCTCTCAGGCTACGCGGTTTGTATGACCCCTACCGAGAG 73048  
QY 3595 ttgctgtgtaaccttaaatgccaatggaagtgttgcgtgctcgaaggttaaaac 3654  
DB 73047 TTGCTGTATTACTTACATGAGCTAATGAGGTGCTGCTTCGTTTGAAGGTAACCTT 72988  
QY 3655 taacaatataacatctgtgct-----cctcctaactccttgacgtggaag 3702  
DB 72987 GGAATTTTAACGTGTTGATCATTAAGTAAGAAAGACATCCACATGATGTATAGTGTG 72928















KEYWORDS	HTG.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE	1 (bases 1 to 65899)
AUTHORS	Lin, X., Kall, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujita, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldmann, T.V., Buell, C.R., Ketchum, K.A., Lee, J.-J., Ronning, C.M., Koo, H., Motil, K.S., Cronin, L.A., Shen, M., VanHaken, S.E., Umayam, L., Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhagen, G.P., Preuss, D., Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C.
TITLE	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
JOURNAL	Nature 402 (6763), 761-768 (1999)
MEDLINE	20083487
REFERENCE	2 (bases 1 to 65899)
AUTHORS	Lin, X.
JOURNAL	Direct Submission
COMMENT	Submitted (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA On Dec 17, 1999 this sequence version replaced g1:4726109.

of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (<http://www.tigr.org/tdb/at/at.html>).

Genes were identified by a combination of three methods: Gene prediction programs including GENAL (<http://arthur.epm.ornl.gov/pub/genal/>), GeneFinder (Phil Green, University of Washington), GenScan (Chris Burge, <http://genome.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curation based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Rddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smil, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to the bottom of the chromosome.

We thank the CSHL/MassU/ABI consortium for sequencing BAC clones F6P23, F5J6, F17A5, and T13J16, the ESSA group for sequencing clone F11304, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khaila, Michael E. Heaney, Lily Fu, Feng Jiang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation Department of Energy and the US Department of Agriculture.

Address all correspondence to: [at@tlgr.org](mailto:at@tlgr.org).

## FEATURES

### source

```
misc_feature      1. 65698
                  /note="Sequence from clone F13J11"
```

gene	CDS
complement(join(.842,907..1234))	complement(join(.842,907..1114))
/gene="At2g13660"	
complement(.1234)	
/gene="At2g13660"	
/note="F13J11.1"	

repeat\_region

```

mRNA      join(<3349. .3541,3644. .3744,4109. .4230,4375. .4471
          4583. .4693,4792. .4928,5629. .5755,5848. .>5997)

```

gene

CDS  
/NOV 1980; FOLIO 3644.  
Folio(3349. .3541,3644. .3744,4109. .4230,4375. .4471,  
4507 4703 4809 5030 5255 5848 5907)

gene

11

gene

cons

rene

repeat\_region

3

repeat\_region



repeat_region	/rpt_family="(CAA)n" complement(31832..31913) /rpt_family="(GAA)n" join(32232..32373,32688..33252..>33473) /gene="At2g13720" 32232..>33473 /note="At2g13720" join(313011,7' 32264..32373,32688..33252..33473) /gene="At2g13720" /note="unknown protein" /codon_start=1 /protein_id="AAD28313.1" /db_xref="GI:4726113" /translation="MMAVLQAERLGS MRCSRMPRGEMVRVVPDGLKSSITV LVPTVVYVSRLTESHKYFTVLNL SVYSTLTNRGKKDDNRYLNRTIKYKGVW OVMEKNVLSVNDOKRFFIVDEGSEAL ENAGTSEDOKQWKIRKYYEGFLEDSKO QMHSDFPTEKKLKLYPTTSNQKHIN INKNTYLKKNDKSKKIILKKYEKKVI QLIIL"
gene	
CDS	
repeat_region	33343..33441 /rpt_family="POLY_A" complement(33782..>34228) /gene="At2g13730" complement(33782..>34228) /gene="At2g13730" /note="P13311.8" complement(33782..34228) /gene="At2g13730" /note="unknown protein" /codon_start=1 /protein_id="AAD28314.1" /db_xref="GI:4726114" /translation="MADRVTRADAKGY MVREDDOAYNEAGRUDHALPPEDAE DQARLNQIALPADAGLIYDLHNRCIG IDRIHLQMNEDAIGVEHNRGVDREHO PQIPAIPL AAEPRTLGDFNRPGMTYANSAIVPP PYRDNYESHLLFYITLH"
repeat_region	complement(33589..35776) /rpt_family="(TAAA)n" 35924..37497 /gene="At2g13740" /note="P13311.9"
repeat_region	pseudo complement(37457..37522) /rpt_family="(GAAA)n" join(<38993..39051,39561..39686,39797..39989, 40095.>40208) /gene="At2g13750" 38993..>40208
mRNA	
gene	
Query Match	9.2%; Score 375.6; DB 50; Length 65899; Best Local Similarity 68.4%; Pred. No. 1.5e-61; Matches 536; Conservative 0; Mismatches 244; Indels 4; Gaps 1;
Db	3298 tatgcagccgaagaacatcagaagtcttcggagcgccaagagtttcttcgsg 3357
Db	64024 tttagcacctaataaaaccacagaggttaatttaaaggacattaaaaggttcccttcggt 64083
Qy	3358 gagctacaagtgscgagtgatggtttaagtaaacagaacattttggcagaagtggttt 3417
Db	64084 GAACGTCTAGTTCTCTACAGGAATTTTACCACAAGAATAATGATTGGCGAAGGACGTTTT 64143
Qy	3418 gggaaagctcaacaaggagcgcttggcgaagcggaaactcttgtctgtctaagaagactaag 3477
Db	64144 GGTAATATTGTAATAAGACGTTTAAGCTGATGACACTCTTAGTGCGTGAAACGGCTTAAT 64203
Qy	3478 gaagagcgaactcaggtgtagagcgcccaggttccaacaagaagtatagatatgatcatg 3537
Db	64204 GAAGAAGCTAACCAAGGTGGGGAACACGACGTTTCACAAACGGAAGTGAATCATCAGTATG 64263
Qy	3538 gcagctcatcagaacctgttagattaagcagaaggtttctgtatagaacgccgacagagatg 3587
Db	64264 GCCCTTATATGGAACCTGCTTCGGCTTCGTGGCTTTGGATCATCCCAACTGGAAGATTA 64323

QY	3558	cttgggtatccttatactgccaatgggaagtgttcttgctgtctcaagggtataaacaa	3657
QY	3658	acaattacaacatctgtgtctctctcaattacttgacgltgaagtgtttttcatgltt	3717
Db	64324	CTTGTATTATCCCTCAATATGGCTAAATGGAAGTGTGGCTTTGTTTAAAGAGTACCAATTC	64383
QY	64384	TTTGTCCTAAAGGTTATTATGTCATATATTTTGTGTGTTTGTGACTGGTTGTATTTGGT	64443
QY	3718	tcctttatgggttcaataatgttggttacactaatgacacagaagagccacgltcaaac	3777
Db	64444	CAGTTGTAAGATCTATCTATCTTTGATT----TTCATATTATAGACGCTCCTGAAGCAATC	64499
QY	3778	cttcgctgtgattggccaacgoggaagaatcgcgtgctgctcagctgtagtggttgtctt	3837
Db	64500	CAGCCCTTGACCTGGCCAAAAGAAAAGCATATTGCTCTGGATCACACAAAGGGGCTGCAT	64559
QY	3838	acctacatgatctacgcgcgcacccaagatcattcacccgltgaactaaagcagaacatcc	3897
Db	64560	ATTTACAGATCATTTGGCAGCCAAAGATCATTCACCTGGATGTGAAGCTGCAATATAC	64619
QY	3898	tcttgaacgaagaatttcgaagcggltgttggaagatttcgggttggcaaatgtagact	3957
Db	64620	TCTTGATGATGAAGATTGGAAGCTGTGTTGGAGATTTGGCTACCAAAATTAATGAATT	64679
QY	3958	ataaagaacactacgtgtaacaacagacgctcggtgacacatgctgtacatgctcagat	4017
Db	64680	ATTACGACCTCCATGTGAACTGCTGTACGGGGTACGATTGTGCCATATATAGCGCCGAGT	64739
QY	4018	atctctcaaccggaaatcttcagaagaacccgaagcttttcgatacagatcatgcttc	4077
Db	64740	ACCTCTCGACAGGAATAATTTGTGAAGAAGCTGATGTTTTSGGTACGGGGTCAATGCTTC	64799
QY	4078	taga 4081	
Db	64800	TCGA 64803	
RESULT	8		
AC002292/c			
LOCUS	AC002292	120787 bp	DNA
DEFINITION			Genomic sequence of Arabidopsis BAC F8A5, complete sequence.
ACCESSION	AC002292		
VERSION	AC002292.1	GI:2252639	
KEYWORDS	HTG.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 120787)		
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R., Au,M., Araujo,R., Brendel,Y., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vysotskaia,V.S., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.		
TITLE	Genomic sequence of Arabidopsis		
JOURNAL	Unpublished (1997)		
REFERENCE	2 (bases 1 to 120787)		
AUTHORS	Federspiel,N.A., Davis,R.W., Conway,A.B., Palm,C.J., Conway,A.R., Kurtz,D.B., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Shinn,P., Sun,H., Oji,O., Osborne,B., Shen,Y.K., Toriumi,M., Vysotskaia,V., Yu,G., Theologis,A. and Ecker,J.		
TITLE	Direct Submission		
JOURNAL	Sequencing (05-JUN-1997) Biochemistry, Stanford University/DNA		
REFERENCE	3 (bases 1 to 120787)		
AUTHORS	Federspiel,N.A., Davis,R.W., Conway,A.B., Palm,C.J., Conway,A.R., Kurtz,D.B., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Shinn,P., Sun,H., Oji,O., Osborne,B., Shen,Y.K., Toriumi,M., Vysotskaia,V., Yu,G., Theologis,A. and Ecker,J.		
TITLE	Direct Submission		



JOURNAL	Submitted (12-JUN-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	4 (bases 1 to 120787)
AUTHORS	Federici, N.A., Palm, C.J., Conway, A.B., Kurtz, D.B., Conway, A.R., Au, M., Araujo, R., Brendel, V., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y., Oj, O., Osborne, B.I., Shinn, P., Sun, H., Toriumi, M., Vyotskala, V., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUL-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	5 (bases 1 to 120787)
AUTHORS	Federici, N.A., Palm, C.J., Conway, A.B., Kurtz, D.B., Conway, A.R., Au, M., Araujo, R., Brendel, V., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y., Oj, O., Osborne, B.I., Shinn, P., Sun, H., Toriumi, M., Vyotskala, V., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE	Direct Submission
JOURNAL	Submitted (11-JUL-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	6 (bases 1 to 120787)
AUTHORS	Federici, N.A., Palm, C.J., Conway, A.B., Kurtz, D.B., Conway, A.R., Au, M., Araujo, R., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y., Oj, O., Osborne, B.I., Shinn, P., Sun, H., Toriumi, M., Vyotskala, V., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE	Direct Submission
JOURNAL	Submitted (02-OCT-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	On Jul 11, 1997 this sequence version replaced g1:2251217.
FEATURES	Location/Qualifiers
source	1. 120787 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" 177..2876 /gene="F8A5.1" join(177..214,632..707,1132..1152,1283..1479,1617..1761,1845..2159,2256..2526,2788..2876) /gene="F8A5.1" /note="Hypothetical protein" /codon_start=1 /protein_id="AAB71954.1" /db_xref="GI:2462735" /translation="MCCRWYTLRQNDGTEVSLVFGSMQVLAIPPIGFFKQIYLQI LNTVLDLVKQGTSPGNOTQYFVAVNADKPPCTASVLSLEKSGLEDVYLRANCLDT MCDYCFDREHLYKSGTVOGDKHAKVGGTMEVYVLSILESTOHCRSRGKREY LYRNPFPSPYPHALASALKKIDPKRKGILLAVNPDGQVFLQEMNFTSYQIQIAL HWINQPTKQNGPGISLAKGKKNALDKAHGEGFLSSSRKISYVDLANSI AGLIFFSTIDLDQGDCLSLGFPQIQEVERDTYVNIQKTIYVIGNERKPKQDEAA PFLFDESETSFEDEVDRENSTHEYDATATKATYKQCLATMIQE" complement(3171..4444) /gene="F8A5.2" complement(join(3171..3389,3471..3605,3706..4029,4418..4444)) /gene="F8A5.2" /note="nearly identical to rice water stress induced protein gp1d26537 537404" /codon_start=1 /protein_id="AAB71970.1" /db_xref="GI:2462751" /translation="MAPEISVNPMTYLSERKHAQAPRRAYVTPIAGNDYKGYGLAK GLRWKSAIVLVAMLPDVEHEHILRSQGCYVARETEVPYPPDNOVERAMALYVNIY SKLRINWEEYSKMIYLDADIQVDNDLHFDLSDAFYVMDCFCQEKWSSHQISI GYCOCEKPEVYKPEDESPPPPLTFNAGMVFEPSPSYLQTLQLEITPPSPFADQ FLNMFEEKVTPIPLVYNLVLAIMRHEVNELEKVVHYCAAGSKPMRYGSEENM DREDIKLVKMDVYNDESIDERSKIPAAEEYVTKSSILASVLEPTEYPPAPSA "
gene	complement(8347..9066) /gene="F8A5.3" complement(join(8347..8443,8539..8657,8941..9066)) /gene="F8A5.3"
CDS	

gene	/codon_start=1 /product="putative ADP-ribosylation factor" /protein_id="AAB71955.1" /db_xref="GI:2462736" /translation="MGISPAKLEFSLFAKEMRIIMVGDAAKTTILYKLGIDE LRVAVLIPANKQDLMNANNAEITDKGLSHLRQHRHYIQITCATTEGLEGIDWL SNNIANKVSFF" complement(9782..14180) /gene="F8A5.4" /note="Atyps34" complement(join(9782..9919,10031..10178,10275..10418,10509..10663,10780..10917,10991..11164,11286..11477,11815..11936,12047..12182,12300..12395,12475..12622,12803..12981,13104..13313,13414..13572,13659..13755,13867..13963,14069..14180)) /gene="F8A5.4" /note="location of EST gb w43464" /codon_start=1 /product="phosphatidylinositol 3-kinase" /protein_id="AAB71971.1" /db_xref="GI:2462752" /translation="MGANERFFLSCDINSVYTFRIEKIDGNLPYKSSDSGVSI AEKPELYTECALYIDAPFGLPWRRLKTKTGPYCMNELITLSKYRDLTAHQALIT VMVYSGKTEGLIGATVLFNFKMKMGKGRRLRMQGEKDGSEFPTSGYKREH RGELETERKIMNKYERGOIOSIDMLRLKSLDITKEOESYKHSGLFYIDPCEP EHRVYPOESGANFTIAPIGSTNEEYVTDTELGRTNPSENKQLTARSIDRGITDR LKPSNTERKISQRYLVKPPRKLTSGRDQRLKRPPLMSERBALTKRCVWESDVO EAKQALDNTKWMIDVCDMLDELSPLESEYRAYAVSLRADDELOCTLDLVO ALRERSDRSCIQFLVQALQNIETASLNNYVAVELDHVARYSTYLEENI IKLPYGVNGDGYQWQSVLRQIETLAQCSITREVRNHYARKIEKRLQGLGL SEULEEPEERISPLTPNVLKLGIVAGESLFAKSLAPLTPRTEEGSGSLIRKQ DDLRODQVVOVWMLMDRLKLENDLCLTPKVAIVAGDEMLEFISRSQAQIIS EHSITSTYAKOEHFDEHAPGRTATCLDTEFKSCAGSVTYTILGIDHLOVLLTD GRLEPHDFAITIGDRPKPPPKTKCKEVAEAGSQYTRRSYCCCEANITRS SNTLINFILMASTTIPDIASDEKSTILQEKFRIMDDEACIFPDLLINESVAL FQWVETIHRMAQYWR" 15572..17812 /gene="F8A5.5" join(15572..16004,16236..17812) /gene="F8A5.5" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71966.1" /db_xref="GI:2462747" /translation="MVKKVAATKNSPSLAIAKKSRSNKDVSYEAPLISYNDRI RPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6" join(18712..19150,19685..19839,19981..20987,21054..21188,21973..22091,22335..22510) /gene="F8A5.6" /note="similar to "Mx" GTP-binding proteins" /codon_start=1 /protein_id="AAB71956.1" /db_xref="GI:2462737" /translation="MGSSKHVYTPRTSSPLAIVQAMPDNRVYVIEAPLISYNDR IRLPLDVTYRLNLMVNRGILPFIYVGOSSGKSVLSLGISLPGQGLCTEVP LVMRLOSSSPSEIPVNDKVPYDDEHIAELIRATVYVINGSGVSDAPLTLHK KATVPDITMDLDEGTRVYVNGCPENITVQISGMNEYREPESITLVLSATVPTT CESITRSKATYDKGQRTLAVYTKADAPGLQKVTADVSIVLGVCRNLTGETY EEAQMOBELFRTHPVLSLDEIVIGIYLAQKLMLOSMTARLPKIVSKINOKLD TAVLEKQKLEPVNASTGEALMALMDITGSKESLAILVQGFSEYPPDOHMCARL ADMISOFSQDILQAKPEVAFLMEDEIKIIDECKCVGLPFIIRSAFLLISQHVQI DKPEYFINKIMDIYEDVLSYTAKRSDNPFOOSIKRGRMLIKIKIOSVARNYI VEMKELTDYCNPEYMTSWTKTQSAOESITLAVVKNENIDPEFVYGGGNYISLKR YAHILTPADMKMRITSYKTYLRRITVNLATYQLSKSYLVNRFOKEIYAEAVDP RDGGVEMKLEESPLVASKREKLQNSIKILKESKQVAAIVDQC" 18712..22510 /gene="F8A5.6"
------	---



```

PEXMTSWTQMTAEQONFIYAVLSDGKKPEHFIATGEGGIVIKLSIKRYHAHLLQOAFD
MKRRIASVYTVIRIVDSIALYLIQAGGLORMLEPSVASKREKXIKNSIKIPKES
KDAVAALIVQSSSELLALSLKGLSTITEELIMDEBILALISAAAPALASCKGV
STPLTLHVKKADVDPDLVLVDPEITRVNQTQNIHAKNIQSDYEVHQAAT"
gene
/ gene="F8A5.7"
join(23787..24210,24676..25157)
/ gene="F8A5.7"
/ note="similar to "Mx" GTP binding proteins"
/ codon_start=1
/ protein_id="AAB71957.1"
/ db_xref="GI:2462738"
/ translation="MGSKSNDEYEDVAGKSSLYNTPIEAPIVASSYNDRIPL
DLYDRILNINMVEGDIOLPIYVNGDSSGSSVLESLAGILPRGQICFPLVLM
LROSSPEPEIWEYSDKVPDDEHVAAICADYVAGEVSDPPLTSYKKN
VPDLIVDLPGLTRVYNQPERITVQISRMKTIKIEPESITLVLSATDPTFCES
IRMSRQVDRKERTLAVTVRKADAPGLQKTVADVSIGLGYICVRNRIGETVEEA
RVQEDLFRTHPLSLIDIDIVGIL"
gene
complement(26476..28975)
/ gene="F8A5.8"
complement(join(26476..26936,27030..28070,28540..28975))
CDS
Query Match 7.1%; Score 289.6; DB 8; Length 120787;
Best Local Similarity 63.0%; Pred. No. 2.3e-45;
Matches 494; Conservative 0; Mismatches 264; Indels 26; Gaps 2;

QY 3298 tatcagccggaagaatccagaagtcacatcctggagacgtccagaaggtttcttcgag 3357
Db 109090 TTTCCAGAACAAATGATGCCAGAGAGTATTGAGGCATTGGAAGAGTATACATTCMAA 109031

QY 3358 gaactacaagtgcgaatgtagtgggttttagtaacaagaacatttggcgaggttggtt 3417
Db 109030 GAGCTTAGATGTGCGACCAACATTTCACTCGAAGAAATTCGGAAGAGCGGATAC 108971

QY 3418 gggaaagtctcaaggagcgttggcgaggaactctgtctgtcacaagactcgaag 3477
Db 108970 GGGATTGTGTGAAGGACACTTAAACGATGGAAGTCTTGCTGTCAACACTCTCAAG 108911

QY 3478 gaagagcgaactccaggtggagagctccagtttcaacagaagaatagatgataagtatg 3537
Db 108910 GACGTGTAACATGGCGGTGGAGAGTCCAGTTTCAGACAGAAAGTAAAGACTATTAAGTTTG 108851

QY 3538 gcaattcactgaaacctgttgaatgaatgaaggtttctgtatgacacgagcagaagattg 3597
Db 108850 GCTCTTCATCCCAATTCCTCCGCTCCGCGTTTCTGTAGTAGCAACAGAGAGAAATT 108791

QY 3598 ctgtgtatccttcaatgagcgaatgagtggtcttcggtcccaaggttaaaactaa 3657
Db 108790 TTAGTCTACCTTACATGCCAAATGGAGGTGTCGATCAGCCTTAAAGGTAATGACAGTT 108731

QY 3658 acaatlaacactctgtgctctcctcaattacttgcagtgaaagtgttttttcattgtt 3717
Db 108730 TCCAAAAAATGCTGAATTTT-----TCAAGCCTTTTGGCAATTT 108689

QY 3718 tccctttaggttcataatgtgtgttaactaaatgacacagagagccacgctcacac 3777
Db 108688 TTAATAGTTTCATC-----TTGCTACATTTTGTGTTAGATTAATATCCGTGAGAGC 108637

QY 3778 ctccgctgattggccaacgcggaagaagaatcgctaaagtcgaagtcgaaggttgcctt 3837
Db 108636 CAGCCTTAGATGTGTCGAGAAAGAAAGATAGCGGTTGGACAGCCAGAGAGACTAATTT 108577

QY 3838 acctacatgactcgtgcgtccgaagatcattcaccgttgacgttaaaagcagcaaacatcc 3897
Db 108576 AACTTACAGAGCAATGTGACCCGAAAGATTATACACCGGATGTGAAGAGCTTAACATTC 108517

QY 3898 tcttgaagaagaattcgaagcgttgttggagattccgggtttgccaagcttaatgact 3957
Db 108516 TGTGTGATGAGGACTTCAACAGATTGTTGGTGTGATTTTGGGTGCTTAAGCTTCTTAAC 108457

QY 3958 ataaagactcaacgtgacacagcagtcgtgacacatcgtgcacatcgtctcagaat 4017
Db 108456 ATAGAGACTCTCATGTCTCAACATGTGACGTCCGTGGAACGTGTGGCCACATTCGACCTAGT 108397

```

```

QY 4018 atctctaacccggaagaatcttcagagaaacccgacggtttcggatcacgaatcatgcttc 4077
Db 108396 ACTTATCCAGCGGGTCAGTCTTCAAGAAAGACTGATGTCTTGGCTTTGGCATCTTCTCC 108337

QY 4078 taga 4081
Db 108336 TTGA 108333

```

```

RESULT 9
AB012245/c
LOCUS
DEFINITION
Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone: MRAl9,
complete sequence.
ACCESSION
AB012245
VERSION
AB012245.1 GI:3046853
KEYWORDS
HTG.
SOURCE
Arabidopsis thaliana (strain:Columbia) DNA, clone:MRAl9.
clone:MRAl9.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
REFERENCE
1 (sites)
Nakamura, Y.
AUTHORS
Nakamura, Y.
TITLE
Structural Analysis of Arabidopsis thaliana Chromosome 5. VI
JOURNAL
Unpublished (1998)
REFERENCE
2 (bases 1 to 86722)
Nakamura, Y.
AUTHORS
Nakamura, Y.
TITLE
Direct Submission
JOURNAL
Submitted (23-MAR-1998) to the DDBJ/EMBL/Genbank databases.
Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of
Gene Structure 2, 1532-3, Yana, Kisarazu, Chiba 292, Japan
(E-mail:ynakamu@kazusa.or.jp, Tel: +81-438-52-3945,
Fax: +81-438-52-3934)
FEATURES
location/Qualifiers
source
1..86722
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MRAl9"
/clone_id="Mitsui P1"
BASE COUNT 28628 a 14796 c 14610 g 28688 t
ORIGIN

```

```

Query Match 6.8%; Score 277.8; DB 7; Length 86722;
Best Local Similarity 62.8%; Pred. No. 4.1e-43;
Matches 484; Conservative 0; Mismatches 252; Indels 35; Gaps 2;

QY 3311 aagatccagaagatcatctcgtgagacagtcgaagaggtttcttcggagagctacaagtgg 3370
Db 72179 AAGACTTACGAATTTGAAATCGGCAATGAAAGAGTTCAGTTTGGGAATTAACACCG 72120

QY 3371 cgagtgatgggttagtgaacaagaacatttgggcagaggtgggttgggaaagtctaca 3430
Db 72119 CAACAAGCAATTTTGTCCAAAGAACATTTTGGGCAAGAGAGGTTTGGAGTGGTTTATA 72060

QY 3431 aggaagcgttggcagacggaactctgtgtgtcgaagacgtgaagaaagaggaactc 3490
Db 72059 AAGGATATCCCAATGGAAGTGTGTCGAGTTAAAGATGGA--AAGATCCGATTT 72003

QY 3491 caggtggagagctccagttcacaagaagtagagatgataagcatgacgttcatcga 3550
Db 72002 ATACAGGAGAACTTCAAGTTCAACCGAAGTAGAGATGATTGCTTAGCTGTTCAACCGTA 71943

QY 3551 acctgtgagattgaaggtttctgtatgacacggaacggaagagatgctgtgtgatactt 3610
Db 71942 ACCTTTTACGCTCTTTGGATTCTGTATGACCCCGGAAGAGAAATGCTGTGTATCCGT 71883

```



Query Match	6.7%	Score 273.2	DB 7	Length 86064
Query	3611	acatgccaatgaagatgctgcttcgctgcacagagtaaaactaacattnaacatc	3670	
Db	71882	ACATGCCAAATGGAACCGTAGCTGATGCTCTGAGAGGTTTCAATTAAACATTTCTTGT	718233	
QY	3671	tctgctctctcctcaattcattcttgacgtgaaagtgttttccatgyltttcctta	3730	
Db	71822	TTTGTTATTTCTTTCGATATATATAGCCGAAAGTTCCTTATAGATY-----	717767	
QY	3731	cataattgtgtttacactaaagaaacagagagccacocgtcacactccgcttatg	3790	
Db	71776	-----ACATGTTTCAGACATTTATGAGAGAAAGCCGTCTTAAGATTG	717353	
QY	3791	gccaacgcggaagaagatcgcgcctagagctcacgctcgaaggttgccttaccatgatca	3850	
Db	71734	GAATCGGAGGATATAGCATTTGCACCTCGGCGCAAGCTGAGAGACTTGTATCTTGCACNGCA	716757	
QY	3851	ctgcgcatcggaaatcatctaccacgtgacgttaaaagcagcaacatcctcttaagaaaga	3910	
Db	71674	ATTCATATCCAAATATTTTACACAGAGAGCTCAAAAGCTCCAAATTTTACTTATGAGAG	716151	
QY	3911	attcgaagcagtttgcgttgagatcttcgggttgccaaagtaatgacatataagaacctca	3970	
Db	71614	CTTGAAGCAATATGTTGGGGAATTTTGGTCTTAGCAAAAGCTTTTAAACAGAGAGTTCCAC	715555	
QY	3971	cgtgacaacagcagtcgtgacacatcgtgtcaatcgcgtccagaatatcttcaaccgg	4030	
Db	71554	TGTCACATACGCGACGTGCGAGAACATGTGACACATCCCTCCAGATACCTTTCCACTGG	714955	
QY	4031	aaatcttcaggaagaaacgcagcttttcggatcgaagatcagttcgttttga	4081	
Db	71494	ACAGTCTCTGAGAGAAACCGATGTTTTCGGATTCGGAGTACTAATCTTGA	714444	
RESULT 10				
LOCUS	AB013395/c			
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MON23			
ACCESSION	AB013395			
VERSION	AB013395.1			
KEYWORDS	HTG.			
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MON23.			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)			
AUTHORS	Nakamura, Y.			
JOURNAL	Structural Analysis of Arabidopsis thaliana Chromosome 5. VI			
REFERENCE	Unpublished (1998)			
AUTHORS	2 (bases 1 to 86064)			
JOURNAL	Nakamura, Y.			
TITLE	Direct Substitution			
ABSTRACT	Submitted (06-MAY-1998) to the DDBJ/EMBL/GenBank databases.			
KEYWORDS	Yasukazu Nakamura, Kazuena DNA Research Institute, Laboratory of Gene Structure 2: 1532-3, Yana, Kisarazu, Chiba 292, Japan (E-mail:lynakam@kazarusa.or.jp, Tel:+81-438-52-3935, Fax:+81-438-52-3934)			
FEATURES	Location/Qualifiers			
SOURCE	1. 86064			
	/organism="Arabidopsis thaliana"			
	/strain="Columbia"			
	/db_xref="taxon:3702"			
	/chromosome="5"			
	/clone="MON23"			
	/clone_lib="Mitsui P1"			
BASE COUNT	27710 a 15313 c 14986 g 28055 t			
ORIGIN				

	Best Local Similarity	61.1%	Pred. No. 3.1e-42:	
	Matches	484;	Conservative	0; Mismatches 268; Indels 40; Gaps 1
QY 3290	aaatctctatgcaccggaagaagatccagaatctcatctcggcacaagctcaagagctttc	3349		
Db 74153	AAAGTTTGATTTCCAGGTGAAGTGGACAGAAAGAGATTGCTTTTGGACACTGTGAAGATTGG	74094		
QY 3350	ctttgcggagagctacaaatgycgagtgatgggtttagtaacaagaacatcttggcagag	3409		
Db 74093	CATGGAGAGAGCTTCACTGTTGGCTACTAGAGATGAGTTTACGTGAAGAAAGATGTTCTCGACAG	74034		
QY 3410	gtgggtttgggaaagtctacaagggagcgttgcagacggaactctgttgcgtgtcaaga	3469		
Db 74033	GAGGCTTTGGGAAGATTTCAAAGAGATTGCTTCGGATGGCAGCACCAAGTGTGTATAAAA	73974		
QY 3470	gactggaagagagcggaatccaggcttgagagctccagtttccaaacgaagttagatga	3529		
Db 73973	GATTGACTGATTTTGAACCTCCAGAGAGAGATGAAGACTTTCCAGAGAGAGATTGAATGA	73914		
QY 3530	taagatcgcgaagttcacgaaacctgtgttgaattacgagagtttctgtatgacacgcgcg	3589		
Db 73913	TAAATGATAGCTGTTCCATAGGAATCTGCTTCGCCCTATAGGCTTTGTGTACACAAACTG	73854		
QY 3590	agagatctgtctgtgtctcttcttcaatgagccaaagagaggttgcgtcgtctcagagta	3649		
Db 73853	AAGCACTTTGGTGATCCTTTCTATGACAGATCAATGAAGTGTGCATNTGCTTAAGAGATGA	73794		
QY 3650	aaaactaaacaattaaacatctgtgctctctctcaattactttgacgtgaagtgcttt	3709		
Db 73793	AAAAAAGACACATGA-----GCTAT 73774			
QY 3710	tcatgttttcccttattgggttcataattgttgttaacctaatgaacaaagagggccacc	3769		
Db 73773	AAACTTTTGAATTCAGAGAGTTTAAGATTTTGATCTTAATCAATGACAGAGATTAAAC	73714		
QY 3770	gtcaaacctccgcgttgatgtggccaacgcggaagaaatccgcgtcaggctcaagctcag	3829		
Db 73713	CGGGGATCCAGTTCTGTGATTGGTTTCAGAGAAACAGATTCGGTTAGTGCAGCAGAGG	73654		
QY 3830	tttgtcttaccatacatgatcactcgatccgaagatcatcaccgtgacgttaaaagcagc	3889		
Db 73653	ACTGGAATATCTTCATGAACATTCGCAACCCGGAAGATCATACACAGAGATGTGAAGAAGCTGC	73594		
QY 3890	aaacatctcttgaacgaagaatctgaacgcgtttgttgagatctcgggttgcgaagct	3949		
Db 73553	AAATGCTTACTGATGAGAACTTTGAAACAGTGGTGTGATTTTGGTTTAACCAAGTT	73534		
QY 3950	aattgacataaagacactcactcgtgacaacagcagctccgttgcacacatcgtgtcacatgc	4009		
Db 73533	GGTAGATGTTAGAAAGCATTAATGTAACTACCTCAGGTCGAGAGAACATAGGCTCATATTGC	73474		
QY 4010	tccgaatatctctcaaccgcgaaaccttcagaagaacacgcgttttcgatatcggaat	4069		
Db 73473	ACCGAAATGATATATCCACAGGAATGTCAGAGAAACCGATGTTTCGGGTACGGGAT	73414		
QY 4070	catgctctaga 4081			
Db 73413	TATGCTTCTGGA 73402			
RESULT 11				
LOCUS DCU93048	1755 bp	mrna	PLN	28-JUN-1997
DEFINITION	Daucus carota somatic embryogenesis receptor-like kinase mRNA,			
ACCESSION	U93048			
VERSION	U93048.1			
KEYWORDS	GI:2224910			
SOURCE				
ORGANISM	Daucus carota			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
	eudicotys; Asterales; euasterids II; Apiales; Apiaceae; Daucus.			



REFERENCE	1 (bases 1 to 1755)
AUTHORS	Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
TITLE	A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos
JOURNAL	Development 124 (10), 2049-2062 (1997)
MEDLINE	97313247
REFERENCE	2 (bases 1 to 1755)
AUTHORS	Schmidt,E.D.L., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-1997) Molecular Biology, Agricultural University of Wageningen, Drenthlaan 3, Wageningen 6703 HA, The Netherlands
FEATURES	location=Qualifiers
source	1..1755
	/organism="Daucus carota"
CDS	/db_xref="taxon:4039"
	34..1695
	/note="SERK"
	/codon_start=1
	/product="somatic embryogenesis receptor-like kinase"
	/protein_id="AAB61708.1"
	/db_xref="GI:2224911"
	/translation="MNRNSIIILMYOTPDATYLDKGYAMTELYSNNGSPISDDG NLTIVLSIDLTMNSFSGPIPTDLKRLTRLRLLRNNSSLSPPIMSLNTITLQYLDDI SNNLSGPVDPNGSFSLFTPISEANNINLCGRPVGRCPGPSPEPFPSTVDLP FGSGPLGALGVGAAGAAALFLFAAPAVAWRRKRKEPHFDVAEDPEVLHGLQIK RSLSRELOVAIDTFSTILRGSGFKVKYGRADSLVAVKRLKEERTGGELQTEVEE EMISMAVHRNLRLRGPCMPETERLLVPYVANGSVASCLERPSSEPLDMPTRRKR ALSGARGSLYHDHCDEPKTIHRDYKANIIILDEFENAVGPGFLARLDYDFHTTA VRGLGTAYIAPEYLSLTGKSSSEKTPDGXGIMLELTIGORAPDIARIANDDVMLDDW KSLRNKEULEVDPDLENNTIDIEVEDLIOLVALICTGSGPMERKASEVYRMLEDGD LAEEWMOEKVHYHODELAPHRTSEIIDSITDNLAHFELSGPR"
BASE COUNT	506 a 347 c 407 g 495 t
ORIGIN	
Query Match	6 6%; Score 270.8; DB: 8; Length 1755;
Best Local Similarity	64.6%; Pred.No.1.5e+41;
Matches 504: Conservative	0; Mismatches 157; Indels 119; Gaps 2.
Oy	3302 cagccgaagaatccaagaattcatctctggcagcgctcaagaaggtttcttcgcyggagc 3361
Dd	647 CAGCGTGAAGAGGACCCAGAAGTGCAACTTGTCGCACTGAAGAAGGTTTCTCGCGAAT 706
Oy	3362 taacaagtgcgagtgatgggttagtaacaaagaacattttggycagaggttggttgga 3421
Dd	707 TCACAAGTCGCCAACGGATTCTTAGT-----ACCATACTTGGAAAGGTGATTGGTA 760
Oy	3422 aaggtcacaaaggagacctggcgacagaggaactcttgtctcaagaagactgaagaag 3481
Dd	761 AGGTGATATAGGAGACCCCTTGCTGATGCGTCACTTTAGCAGTTAAAAGCTTTAAAGAA 820
Oy	3482 agcgaaactccaggttggaagctccagttccaacagaagtagatgatattgycag 3541
Dd	821 AACGAACACCAAGTGGTGAGCTGCATTCCAACAGAAGGTGAAAAAGATTAGATGGCTG 880
Oy	3542 ttcatcgaaacctggtgaattaccagaggtttctgtatgcacacgaccggagatctgtg 3601
Dd	881 TCGATGGAATCTTCGCGCTTACGIGGTTTTGCAATGACCAACAGAGCGGCTTCTTG 940
Oy	3602 tgatacttacatgscaatggaagtgttgctcgctccaggtctcaagagtaaanaactaa 3661
Dd	941 TATATCATATACAGCGCTAATGGAAGTGTGCGTGGTTTAAAG----- 984
Oy	3662 ttaaacatcttgctctctctcataattcttgacgtgaagtgttttttccatgtttct 3721
Dd	984 ----- 984
Oy	3722 ttatgggtcataatatgttggttacactaaagacagagagcaccggtcacacactcc 3781
Dd	984 -----AGACCGTCAAGCCATCAACAACATCC 1007
Oy	3782 gcttgatggcaacgcggaagaacatcgcgctagctcagctccagaggttggttctaact 3841

QY	1008	CGTTGATGGCCACACTAGGAAGAGGAGATTGCACATAGATCTGCTAGAGGGCGCTTTCTATT	1067
QY	3842	acatgatcatcgtgatccgaagatcatcccgtagcgttaagaagcgaacatctctt	3901
Db	1068	GCATGACCAATTGATGCATCCCAAGATTATCCATCGGATGATAAAGCTCCAAATATATT	1127
QY	3902	agacgaagaattcgaaacggttgcttgagaattcgggtgttgccaagcctaattgactaa	3961
Db	1128	GGACGAGAAATTGAGGCGCTTTGAGTGATTTTGGTTAGCTAGCGCTAAGGATTACAA	1187
QY	3962	agacactcacgttgaacaacagcagtcgtagccatcatcgtgtcacatcgtccagaatact	4021
Db	1188	GGATACCCATGATTACCACTCTGTAAAGGGGTACTTTGGCGTACATACCTCCGAGTACT	1247
QY	4022	ctcaacgcggaatccttcagagaanaacgcgcttttcggataggaatcatcgtcttga	4081
Db	1248	CTGCACTGGAAATCATCATCAGAGAACCGCATGCTCTTTGGTTATGGGATTATGCTTTAGA	1307
RESULT	12		
LOCUS	A67797	1815 bp	DNA
DEFINITION	Sequence 2 from Patent WO9743427.		PAT
ACCESSION	A67797		
VERSION	A67797.1	GI:4756623	
KEYWORDS	.		
SOURCE	carrot.		
ORGANISM	Daucus carota		
REFERENCE	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids II: Apiales: Apiales: Daucus.		
AUTHORS	1 (bases 1 to 1815)		
TITLE	De,V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.		
JOURNAL	PRODUCTION OF APOMICRITIC SEED		
FEATURES	Patent: WO 9743427-A 20-NOV-1997;		
source	CTBA GEIGY AG (CH)		
	Location/Qualifiers		
	1..1815		
	/organism="Daucus carota"		
	/db_xref="taxon:4039"		
	94..1755		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAB42241.1"		
	/db_xref="GI:4756624"		
	/translation="MNRNSINILNYMOTDAYLDKXGVLMTLEYSNNISGPIPSDGL		
	NLNLIVSLIDLYMNSFGPIDPTLGKILRLFLRLNNNSLGPIDMSLITNTIIVLIDL		
	SNRSLSPVDNGSESLFTPISEFANNMLNCGPVGRCPSPSPSPPIPISTVOP		
	PGONGPFGATAGGAAAGAAALFLPAAAPAAAFMRMRKRREHFPDPAEDEYHGLG		
	RFSIRELQVATDPTPSTTLGRGGGKYYKGLAGCSLAYAVRKREPRGSELROFQV		
	EMISVAHRLNLRIGFCMTPTERLVLVTPMANGSVASCI.REROPSPDLPFRKRI		
	ALGSARLSTLHDHCDPKI.IHRDVKANILLDEEFVVDGFLARLMDYKDVHTTA		
	VRGLTGLIAYEYSTGKSEKSTDFVFGIMLELITQRAFDLARLNDVDYMLDMV		
	KSILKEKLEMLVDPDLENNYIDTEVEVLQVALICTGSPMERPKMSEYVRMLEDGG		
	LAEKWDEMOKVEVHODVELAPHRITSMIIDSITDNLHAFELSGPR"		
BASE COUNT	531 a	354 c	415 g 515 t
ORIGIN			
Query Match	6.6%;	Score 270.8;	DB 5; Length 1815;
Best Local Similarity	64.6%;	Pred. No. 1.5e-41;	
Matches 504; Conservative	0;	Mismatches 157;	Indels 119; Gaps 2;
QY	3302	cagccgaagagaagatccaaagtcatctcgtggagcagcctaagaagatgttcttcgtggagc	3361
Db	707	CAGCTGAAGAGACCCAAAGATGCACTTGGTCACTAAGAGGTTTCTCTCGCAGAT	766
QY	3362	tacaagtgcgagatgatggtttaagaacaagaacatttggcgagagtggttgga	3421
Db	767	TGCAGTGGCAACGGATACTTTAGT-----ACATACTTGGAGAGAGTGGATTGGTA	820
QY	3422	aagttacaaaggagcgtcttgcaagacggaacatcttgtttgctgtcaagaagactgaagag	3481



```

Db      821 AGGTGTATAGAGACGCTTGTGCTGATGCTCATCTGTATGACATTAAGGCTTAAAGAG 880
QY      3482 agcgaactccaggtgagagagccacgtttcaacaagaagtagagataagatgagcag 3541
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      881 AAGCAACACCAGGTGTGAGCTGCACTTTCACACAGAGTGAATGATTTAGCATGCGCTG 940
QY      3542 ttcactgaacacctgtttagattacagaggttctctgtatgacacagcagagagatgcttg 3601
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      941 TGCATGGAATCTTCTGCGCTACGCGTTCCTGATGACACACAGAGCGGCTTCTTG 1000
QY      3602 tgcattccacttagcgcacatggaagtgtgtctgtctgtctcagaggtataaaactaaacaa 3661
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1001 TATATCCATACATGCTTATGGAAGTGTTCCTGCTGCTTTAAAG----- 1044
QY      3662 ttaaacatctgtgtctctctcctaacttcttgacgtgaagtgttttctatgttctct 3721
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1044 ----- 1044
QY      3722 ttatggttcataatgttgtgttaactaatagacacagagagccacgttcaaacctcc 3781
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1044 ----- 1044
QY      3782 gcttattgccaacgcggaagaagaatcgcgttagcgtcagctcgaaggttgccttaact 3841
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1068 CTTGATTTGGCCAACTAGGAAGAGATGTGCTACTAGGATGTGCTAGGGGCTTTCTTATTT 1127
QY      3842 acattgacactgcgtacgaagaatcatctacacgtgtacgttaaaagacagaacatctctc 3901
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1128 GCATGACCATTTGATGCCACAGATTTTCATGCTGATGTAAGGTCGAAATATTTTAT 1187
QY      3902 agacggaagaattcgaacgcgtgtgtgagattcgtgtggaagcgaacgaattgactataa 3961
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1188 GAGCAAGATTTGAGGCTGTCTGTAGTATTTGGGTTAGCTAGGCTCATAGATTAACAA 1247
QY      3962 agacactcagtgacaacagcagctccgtgacacatcagtcacatgcctcagaatctc 4021
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1248 GGAATACCAATGTACAACTGCTGTAAAGGGCTTACCTGAGCAATGCTCCGAGTACT 1307
QY      4022 ctaaacgcggaatactctcagagaagaacgcgttttcgatatcgaatcgtctctaga 4081
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1308 CTGACTGGAAAGTCATCAGAGAAAGACCAATGTCCTTTGGATTATGCTCTTAA 1367

```

**RESULT 13**  
**AB025639/c**  
**LOCUS** Arabidopsis thaliana genomic DNA, chromosome 3, pl clone: MML2, complete sequence.  
**ACCESSION** AB025639  
**VERSION** AB025639.1 GI:4589445  
**KEYWORDS** HTG.  
**SOURCE** Arabidopsis thaliana (strain: Columbia) DNA, clone: MML2, clone: MML2.  
**ORGANISM** Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
**REFERENCE** 1 (sites)  
**AUTHORS** Nakamura, Y.  
**TITLE** Structural Analysis of Arabidopsis thaliana Chromosome 3. II  
**JOURNAL** Unpublished (1999)  
**REFERENCE** 2 (bases 1 to 84896)  
**AUTHORS** Nakamura, Y.  
**TITLE** Direct Submision  
 Submitted (02-APR-1999) to the DDBJ/EMBL/GenBank databases.  
 Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2, 1532-3, Yana, Kisarazu, Chiba 292, Japan  
 (E-mail: ynakamura@kazusa.or.jp, Tel.: +81-438-52-3935, Fax: +81-438-52-3934)  
**FEATURES**  
 source  
 1. 84896  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"

```

BASE COUNT      27260 a 15033 c 15617 g 26980 t
ORIGIN
Query Match      6.4%; Score 263.2; DB 7; Length 84896;
Best Local Similarity 61.6%; Pred. No. 2.4e-40;
Matches 483; Conservative 0; Mismatches 253; Indels 48; Gaps 2;

/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MML2"
/clone_11b="Mitsui P1"
Query Match      6.4%; Score 263.2; DB 7; Length 84896;
Best Local Similarity 61.6%; Pred. No. 2.4e-40;
Matches 483; Conservative 0; Mismatches 253; Indels 48; Gaps 2;

Db      3298 tatgcagccgaagaagatccagaagttcatctggaacagctcaagaggttcttcgcg 3357
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67679 TACACAGACCAAAACAGAGAGAAATGTGTAGGAACTTACAGAGATTATTTCAAA 67620
QY      3358 gagctcaagtgaggagtgatggtgttagtaacagaacatttggcagaggtgggtt 3417
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67619 GAACTTCAATCCGCACTGTATCTACAGACAGCAAGAAATCTGTCGGAAGAGAGGCTTT 67560
QY      3418 gggaaagttcaagaagagccttgcagacaggaactctgtgtcgttcaagaagactga 3477
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67559 GGAATGTGTATTAAGGTGTCTTCATGATGGAAGTATCATCGCGTGAAGATTAAG 67500
QY      3478 gaagagcgaactccaggtgtggaagctcagatttcaacaagaagtagatgataatg 3537
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67499 GATTTAAACAAATGTGTGAGAGAGGTTCAGTTCAGACAGCTTGAATATATAGCCTT 67440
QY      3538 gcagttcacagaacacctgtgtgattacagaggttctgtatgacacagccagagatg 3597
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67439 GCGGTCCACGGAGATCTCCGCTTATACGCTTCTGTACTACTCTCTGAACGGCTT 67380
QY      3598 ctgtgtatccttcaatgagccaaatggaagtgttctgtgtcctcagaaggttaaaactaa 3657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67379 CTGTTATCTCTTACATGTCATGACATGACATGCGCTTCTCTCCTCAAGATTAATTA 67320
QY      3658 acaattaaacatctgtgtctctcctaacttcttgcgtgaaggtgttttctatgtt 3717
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67319 AACA-----CCTTGATTTTGAATTTCTTACTAGTAATGACCCGTT 67279
QY      3718 tccattatggttcataatgttgtgttaactaatgacacagagagccacgcctcaaac 3777
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67278 TACGATTCGTGTCAATGATGTTGCACCTA-----AAC 67248
QY      3778 ctccgcttattggccaacgcggaagagaatcgcgttagctcagctcagaggttgcctt 3837
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67247 CGGTATTGGATTGAGGCAACAAAGACCAATGACATTAGAGACAGAAAGAGGTTGCTGT 67188
QY      3838 accatcatgacatctgcgtccgaagatcattcaccgtgtacgttaaaagcagaacacac 3897
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67187 ATTGCATGAGCAATGTGATCCAAAGATCATTCACCGATGATCCAAAGCTCGAATATAC 67128
QY      3898 tcttaagcgaagaattcgaagcgtgtgtgtgagatttgggttggcagaagtaatgact 3957
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67127 TTCTTACGATTAATCTTTGAGCTGTGCGAGATTTTGGGTTGCTTAAGCTTTTGATTC 67068
QY      3958 ataaagacactcagtgacaacagcagctcgtgtgcacacatcgttcacatcgcctcagat 4017
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67067 ATAGAGTGTCGATGTGACACACCGCGGTGAGAGAAAGTGGGTGACATTCACACTGAGT 67008
QY      4018 atcttcaaccggaataatctcagagaaacagcagcttctggatagcgaatcatgtctc 4077
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67007 ATCTTCAACAGCAATCTTGTGAAAGACAGATGTGTGCTTGGGATTTCTTCTTC 66948
QY      4078 taga 4081
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      66947 TCGA 66944

```

**RESULT 14**  
**AC005170/c**  
**LOCUS** AC005170 110211 bp DNA PLN 17-DEC-1999



DEFINITION Arabidopsis thaliana chromosome II section 135 of 255 of the complete sequence.  
ACCESSION AC005170 AE002093  
VERSION AC005170.2 GI:6598448  
KEYWORDS HTG.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 110211)  
AUTHORS Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldblum, T.V., Buell, C.R., Ketchum, K.A., Lee, J.U., Ronning, C.M., Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., VanAken, S.E., Unayam, L., Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D., Nieman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C.  
TITLE Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana  
JOURNAL Nature 402 (6763), 761-768 (1999)  
MEDLINE 20083487  
REFERENCE 2 (bases 1 to 110211)  
AUTHORS Lin, X.  
TITLE Direct Submission  
SUBMITTED (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
COMMENT On Dec 17, 1999 this sequence version replaced gi:3738313. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (<http://www.tigr.org/tldb/at/at.html>).

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://arthur.epm.ornl.gov/pub/graal/>), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://genomic.stanford.edu/GENSCANW.html>), and NetPlantgene (<http://www.cds.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RX/repeatmasker.html>). Genes are numbered from the top to the bottom of the chromosome.

We thank the CSHL/Mashu/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Likin Zhou, Hanif Khailak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: [atetigr.org](mailto:atetigr.org).  
Location/Qualifiers  
1. 110211  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"

/db\_xref="taxon:3702"  
/chromosome="2"  
complement(1..2793)  
/note="Sequence from clone F27L4"  
complement(356..401)  
/rpt\_family="(TAA)n"  
complement(join(1549..2058,2731..3012))  
/gene="At2g23810"  
complement(1549..3012)  
/gene="At2g23810"  
/note="F27L4.1"  
complement(join(1744..2058,2731..3003))  
/gene="At2g23810"  
/codon\_start=1  
/product="similar to senescence-associated protein"  
/protein\_id="AtF1811.1"  
/db\_xref="GI:6598449"  
/translation="MFLILFLVCITFAFVVTNKGAGEALTEGRYREYKIGDYSTWL  
QRRVNGKMKWIKRSLVESKYSKLEAFVNVPSNFKYKHLTALSGCKPDECG  
FHYVPTWTKNTGTHTNPDQTMNAKKEKCPGCGCKAGLLDNVKSAMKVAIVN  
IYFLVFLIVSYVGCACFRNNKRDSYRITGYRP"  
2794..110211  
/note="Sequence from clone T29E15"  
complement(3567..3677)  
/rpt\_family="(TAA)n"  
complement(3570..3641)  
/rpt\_family="(TAAA)n"  
complement(4565..4612)  
/rpt\_family="(CAA)n"  
6379..6484  
/rpt\_family="(CAAAA)n"  
join(<6480..6749,7602..7726,7811..7848,7929..8039,  
8152..8268,8487..8786)  
/gene="At2g23820"  
<6480..8786  
/gene="At2g23820"  
/note="T29E15.2"  
join(6480..6749,7602..7726,7811..7848,7929..8039,  
8152..8268,8487..8557)  
/gene="At2g23820"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAC63656.1"  
/db\_xref="GI:3738315"  
/translation="MKNVNVVFCKSLSLVFLVPSFTSRSHIRFYSAAGASPNAIR  
CAASDPOSQDSVSPNVAAPVSSSSSAIDFLSCTRLKTPRAGWKRDY  
KQPSIADHMYRMGIMALISDIPQVNDKCKMAIVDAIVGIDITPGGSKKE  
KNRRESFALEHMKLIGSGERAKETAEIMREIENSFPAKYVDFDVELLQALEY  
EQKFPQNTGKMASETVSRKKQH"  
complement(6654..6712)  
/rpt\_family="(GAA)n"  
join(<9531..9582,9709..9779,9866..9947,10054..10169,  
10268..>10396)  
/gene="At2g23830"  
<9531..>10396  
/gene="At2g23830"  
/note="T29E15.3"  
join(9531..9582,9709..9779,9866..9947,10054..10169,  
10268..10396)  
/gene="At2g23830"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAC63657.1"  
/db\_xref="GI:3738316"  
/translation="MSNNELLETPEMYLQFPPELKKOMSCSLYLTNKTDNNVAFVKY  
TNRNNSCVBNYGLILPKSKCYLVYMOAKVPSDMOSFEKFIQSYVLAASGVAKKE  
VTRMNSKSGSHVEETKLRVYVCSITTNITSSPTRNGFIQGC"  
complement(11332..11371)  
/rpt\_family="(TAA)n"  
15801..15917  
/rpt\_family="(TAAA)n"  
15853..15918  
/rpt\_family="(CAAAA)n"



Query	Match	Score	DB	Length
54252	GTTCGAAAGATATCTCTGGCTGGTGGTGGTAAATGCTTACACAGGAAATAATTCGGGG	54193		
3386	gtacaagaacacttttggcagagagtggttgggaagcctcaaggaagccttggcag	3445		
54312	GGTTGGGAATCATTAAGAGCTTCAATTCAGGGAACCTTCATGTAGCTACGATGCTTTTA	54253		
3326	atctggacaagctcgaagggcttctcttgcggagagctacaagtgcagtgatggttca	3385		
54372	GTGTCTAAGATATTTTGTGGTCTTATTTGTAAGCAAGCAAGGAGGAAGGTTACTTG	54313		
3209	ttggacaacttcttaataaatacaatcaatcaatgaatcatgttctttaaattcaaacctc	3268		
54432	TTTCGATATAGTGGTTAAGTCAATAGAGAGTAAGCTTATGCTTATGATTTGTTT	54373		
3269	tttgaataaactctgatctgcaaatctctatctgacgacggaagaagctcag---aagtc	3325		
3209	tttgaacaacttcttaataaatacaatcaatcaatgaatcatgttctttaaattcaaacctc	3268		
54432	TTTCGATATAGTGGTTAAGTCAATAGAGAGTAAGCTTATGCTTATGATTTGTTT	54373		
3269	tttgaataaactctgatctgcaaatctctatctgacgacggaagaagctcag---aagtc	3325		
3326	atctggacaagctcgaagggcttctcttgcggagagctacaagtgcagtgatggttca	3385		
54312	GGTTGGGAATCATTAAGAGCTTCAATTCAGGGAACCTTCATGTAGCTACGATGCTTTTA	54253		
3386	gtacaagaacacttttggcagagagtggttgggaagcctcaaggaagccttggcag	3445		
54252	GTTCGAAAGATATCTCTGGCTGGTGGTGGTAAATGCTTACACAGGAAATAATTCGGGG	54193		

QY	3446	accgaactctgtgtgtctgcacagaagcctctaaagaagcgcacccacagctggagagctcc	3505
Db	54192	ATGGGACAGTGGTTCGACAGTGAACCACTTTAAAGATGTGAATGGAACCTCCGGGAATCAC	54133
QY	3506	agtttcaacagagaagtagaagatgataagtaatgacagcttcacatcgaaacctgtgtgattac	3565
Db	54132	AGTTTCTTACTGTAGCTGTGAGATGATCAGCTTGTAGCTTTCATGGAATTTGCTTGCTTAA	54073
QY	3566	gaggttcctctatgacaacccagcaggaagctgtctgtatcttaccctacatggcgaatgaa	3625
Db	54072	TCGGTTATTTCGGAGTTCTACGCGAAGACTCTTGTGTTTACCCTTACTATGTCGAATGGCA	54013
QY	3626	gtgtgtcttcgtgtctcagaggttaaaaaactaaacaataaactctgtctctctca	3685
Db	54012	GCGGCGCCTTACGCTCAAGGACCAAGGAGTCAAAAAAATGGTTTCGATTAACGCTTTAGTAA	53953
QY	3686	atctacttgcgcggaagtggttttcaatgtttcccttaatgatggttcataatgttgta	3745
Db	53952	AAAA---AAATATGAATCACTTGAACTATTTTGTGTAAGTTCAC-----	53908
QY	3746	cactaatgacacagagaagccaccctgcacaaacctccgctgtatgtgcacaacgcggaagag	3805
Db	53908	-----GCTAAGCAGCGTGTGACCTGGAACACCAAGAGAA	53874
QY	3806	aatcgcgctatgctcagctccgaggtttgtcttactactatgatactcctcgtccgaagat	3865
Db	53873	GATGCGCATTTGGAGCTGCACAGAGGGTGTGTTATATCTACACGACGCAATCGCATCCCAAGAT	53814
QY	3866	catcacccgycagctgaaagcagacaacacatccctcttagacgaagaattcgaagcgttgt	3925
Db	53813	TATTACCGAGATGTCAGGACAGCAAAATCTCTCCATGATGATTTTGAAGCAGTTGT	53754
QY	3926	tggagatctcgggttggcagaagtaatgtaataagacatccatcgycgaacaagcagt	3985
Db	53753	TGGGGATTTTGGACTGTGCAAAAGCTACTCAACCAAGGAGTTTACATGTTCACAAACCGCGGT	53694
QY	3986	ccgctgcagaccatcggtacacatcgtccagaatactctcaaacccggaatacttcagaagaa	4045
Db	53693	TAGAGGAATCTGTGTGCACACTTGCACCTGATGTCTCCACCGGTCAATCATCGAAGAA	53634
QY	4046	aaccgacgctttcggatagcgaatacatgcttctaga	4081
Db	53633	AACCGATGCTTTGGGTCCGTACTCTTTTGCTAGA	53598
RESULT 15			
LOCUS	AB005234	89779 bp	DNA
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone: MBK5, complete sequence.		
ACCESSION	AB005234		
VERSION	AB005234.1	GI:2264306	
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_1lb:Mitsui P1 clone:MBK5.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (sites)		
AUTHORS	Sato,S., Kotani,H., Nakamura,Y., Kaneko,T., Asanizu,E., Fukami,M.,...		
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned pl clones		
JOURNAL	DNA Res. 4 (3), 215-230 (1997)		
MEDLINE	97471969		
REFERENCE	2 (bases 1 to 89779)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUL-1997) to the DDBJ/EMBL/Genbank databases.		



Yasukazu Nakamura, Kazusa DNA Research Institute, Laboratory of  
Gene Structure 2, 1532-3, Yana, Kisarazu, Chiba 292, Japan  
(E-mail: ynakamura@kazusa.or.jp, Tel: +81-438-52-3935,  
Fax: +81-438-52-3934)

## FEATURES

Source  
Location/Qualifiers  
1. 89779  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="MBK5"  
/clone\_id="Mitsui P1"

BASE COUNT 28415 a 16340 c 16159 g 28865 t  
ORIGIN

Query Match 6.0%; Score 244.6; DB 7; Length 89779;

Best Local Similarity 59.8%; Pred. No. 7,66-37;

Matches 468; Conservative 0; Mismatches 279; Indels 36; Gaps 2;

```
QY 3299 atgcagccggaagaatccagaagatcattcgtgacagctcaagaggtttcttgcggg 3358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54502 ATGCAAGGGAAGATGACGAGAAATTCCTTGGACAACATAACGATTCCTTACGTG 54561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3359 agctacaagtggcagtgtagtggttagtaacaagaacatttgggcagaggctgg 3418
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54562 AATCCAGCTCCGACACGATATTTCACAGAGGCAATTTGATAGACAAAGAGATTTC 54621
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3419 ggaaggtctacaaggagcgttggcagacgaaactcttctgtctcaagagactgaag 3478
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54622 GTAAAGTATACAGAGGTTTGCTCCAGCAAAACAAAGTTGACGTGAACGCTTGCGG 54681
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3479 aagagcgaaactccaggttggagctccaggttcaaacagaagtagagataagatag 3538
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54682 ATTACTTCAGTCTTGAGAGAGAGAGCTGTTCCAAAGAGAGATTCAGCTCATAAGCGTTG 54741
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3539 cagttcatcgaacctgttgagattcgaagttctgtatgaaccgcagagagattgc 3598
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54742 CGGTTCATATAATCTTACGCTTATGGCTTGACAACTTCCTCGAAGAAATCC 54801
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3599 ttgtgtatccttaccatggccaatggaagtgttgcgttcgttcagaggttaaaactaa 3658
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54802 TTGTTATCCATACATGAAATCTTAGTGTGACATATCGACTAAGAGGTATATA---- 54857
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3659 caatlaaacatcttgcctcctcaattacttgacgtgaagtttttcaatgttt 3718
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54857 -----TACTTAGCTTCTCTATACATTACCAATAATAGATTCTATATGTGTG 54907
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3719 ccttaagggtcataaattgtgtacataatgacacagaagggccacgctcacacc 3778
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54908 TACATATGTG-----ACATTGTAGATTTGAAAGCGGAGAGGA 54945
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3779 tcgcctgttgccaaagcgaagaatcgcgtaagctcgaagctcgaagttgtctta 3838
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 54946 AGGATTAGACTGGCCAAAGAGCGGTAGCGTTTGGTTCAGCTCACGCTTGAAGTA 55005
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 3839 cctacatgactcgaatccgaagatcattcacgctgaacgttaaaagcgaacaacacct 3898
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55006 TCTACACGAACTGTATACCCGAGATCATACCGCGATCTCAAGCTGCAAACTACT 55065
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3899 cttagacgaagaatcgaagcgggtgttggaagttcgggttgccaagctaatgacta 3958
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55066 TTTAGCAACAATTTTGAGCGCTTCTTGAGATTTCGGTTTAGCTAAGCTTGTGACAC 55125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3959 taaagcactcagctgacaaagcagtcgttgccaccatcggtcacatcgctccagaata 4018
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55126 ATCTCTGACTCATGTCAACAATTCGAGGACCAATGGTCAATTTGCGCCAGAGTA 55185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4019 tctctcaacccggaanaatcttcagagaanaaccgacgttttcggaatacgaatcgttct 4078
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55186 TCTCTGCACAGAAATCATCTGAATAAACCGATGTTTGGTTACGATTAACGCTTCT 55245
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4079 aga 4081
```

Db 55246 TGA 55248

Search completed: June 24, 2000, 00:12:15  
Job time: 45690 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:49:30 ; Search time 446.21 seconds

(without alignments)  
2288.236 Million cell updates/sec

Title: US-09-180-798-20

Perfect score: 4081

Sequence: 1 tctagaacatttgatcat.....tacggacatcgtctctaga 4081

Scoring table: IDENTITY\_NMC  
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4081	100.0	4081	1	V06585 Arabidopsis thalia
2	539.2	13.2	2089	1	V06591 Arabidopsis thalia
3	387.4	9.5	6695	1	V06570 Daucus carota SERK
4	270.8	6.6	1814	1	V06571 Daucus carota SERK
5	108	2.6	3176	1	T62124 Arabidopsis thalia
6	106.2	2.6	9295	1	T62125 Arabidopsis thalia
7	87.6	2.1	4104	1	X07356 Arabidopsis sterol
8	82.8	2.0	3045	1	X23531 Maize Xa21 gene DT
9	82.2	2.0	6644	1	X23531 Maize Xa21 gene DT
10	82.2	2.0	7372	1	X23531 Maize Xa21 gene DT
11	82.2	2.0	7797	1	X23531 Maize Xa21 gene DT
12	82.2	2.0	7996	1	X23531 Maize Xa21 gene DT
13	78.8	1.9	3842	1	X23533 Arabidopsis thalia
14	72.4	1.8	981	1	V06587 Arabidopsis thalia
15	72.4	1.8	1106	1	V06586 Arabidopsis thalia
16	71.6	1.8	894	1	V06589 Arabidopsis thalia
17	71.6	1.8	1554	1	T31307 Arabidopsis thalia
18	71.6	1.8	3293	1	X23532 Tomato Xa21 clone
19	70.8	1.7	788	1	V06588 Arabidopsis thalia
20	70.8	1.7	1063	1	V06590 Arabidopsis thalia
21	64.6	1.6	960	1	T35827 Protein kinase gen
22	64	1.6	5733	1	X00477 Arabidopsis thalia
23	61.2	1.5	2640	1	V22790 Clas I S-receptor
24	59.8	1.5	2443	1	T40344 Tomato Pto protein
25	59	1.4	2685	1	T13425 SRK-A10 gene mutan
26	59	1.4	2686	1	T13425 SRK-A10 gene mutan
27	58.8	1.4	3573	1	T06307 Partial tomato pat
28	58.8	1.4	6471	1	T06306 Tomato pathogen re
29	57.4	1.4	2833	1	O32648 S receptor kinase
30	57	1.4	2704	1	V22792 Clas II S-receptor
31	55.4	1.4	2704	1	V22792 Clas II S-receptor
32	54.2	1.3	960	1	T35826 Protein kinase gen
33	54.2	1.3	2749	1	O63492 S-locus receptor (
34	54.2	1.3	2780	1	V32925 Brassica sp. S-rec

## ALIGNMENTS

35	53.2	1.3	19124	1	T72882	Plasmodium var-7 g
36	51.6	1.3	26811	1	X20253	Borrelia burgdorfe
37	51.2	1.3	684	1	X20377	Borrelia burgdorfe
38	51.2	1.3	3979	1	V14518	CF-5 pathogen resi
39	51.2	1.3	3979	1	V14519	CF-5 pathogen resi
40	51.2	1.3	4123	1	V14523	CF-5 pathogen resi
41	50.8	1.2	9789	1	T41852	CDNA encoding Plas
42	50.2	1.2	110000	1	X20248_08	Continuation (9 of
43	49.8	1.2	1269	1	O47060	ZmpkI homologue pr
44	49.2	1.2	19639	1	X23524	O. longistaminata
45	47.8	1.2	240	1	T76782	Staphylococcus aur
RESULT 1						
V06585						
ID	V06585	standard; DNA; 4081 BP.				
AC	V06585:					
DT	03-AUG-1998	(first entry)				
DE	Arabidopsis thaliana SERK gene.					
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;					
KW	plant breeding; ds.					
OS	Arabidopsis thaliana.					
FH	Key	Location/Qualifiers				
FT	CDS	366..6620				
FT		/*tag= a				
FT		/note= "contains introns"				
FT	Intron	3731..3802				
FT		/*tag= b				
FT		/number= 1				
FT	Intron	3851..3979				
FT		/*tag= c				
FT		/number= 2				
FT	Intron	4124..4211				
FT		/*tag= d				
FT		/number= 3				
FT	Intron	4284..4357				
FT		/*tag= e				
FT		/number= 4				
FT	Intron	4430..4528				
FT		/*tag= f				
FT		/number= 5				
FT	Intron	4642..4757				
FT		/*tag= g				
FT		/number= 6				
FT	Intron	4890..4967				
FT		/*tag= h				
FT		/number= 7				
FT	Intron	5295..5803				
FT		/*tag= i				
FT		/number= 8				
FT	Intron	6197..6339				
FT		/*tag= j				
FT		/number= 9				
PN	WO9743427-A1.					
PD	20-NOV-1997.					
PD	13-MAY-1997: E02443.					
PR	14-MAY-1996: GB-010044.					
PA	(NOVS ) NOVARTIS AG.					
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;					
DR	WPI: 98-086529/08.					
DR	P-PeDB: W47017.					
PS	Production of apomictic seeds - useful in plant breeding					
PS	Claim 26: Pages 64-67; 123pp: English.					
CC	The sequence is that encoding SERK, a putative receptor kinase.					
CC	comprising: (a) transforming plant material with a nucleotide					
CC	sequence encoding a protein which in active form in a cell or					
CC	cell membrane renders the cell embryogenic; (b) regenerating					
CC	the transformed material into plants or carpel-containing					
CC	plant parts; and (c) expressing the sequence in the vicinity					



CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 4081 BP; 1120 A; 770 C; 785 G; 1406 T;

Query Match 100.0%; Score 4081; DB 1; Length 4081;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctagaacaccttctgataagaaataagagtcctcacacacatgggtatgacat 60  
DB 1 tctagaacaccttctgataagaaataagagtcctcacacacatgggtatgacat 60  
QY 61 aatgctgatacttaagaggtacaacatgtaacgcttttatttactttaccct 120  
DB 61 aatgctgatacttaagaggtacaacatgtaacgcttttatttactttaccct 120  
QY 121 actcaaatgtatggcagctttttttttttttaaataagaacagatctgttaa 180  
DB 121 actcaaatgtatggcagctttttttttttttaaataagaacagatctgttaa 180  
QY 121 actcaaatgtatggcagctttttttttttttaaataagaacagatctgttaa 180  
DB 121 actcaaatgtatggcagctttttttttttttaaataagaacagatctgttaa 180  
QY 181 tggatgtgatagaacagtagtaagatcatctgcgcacgcatcactcctccacatg 240  
DB 181 tggatgtgatagaacagtagtaagatcatctgcgcacgcatcactcctccacatg 240  
QY 181 tggatgtgatagaacagtagtaagatcatctgcgcacgcatcactcctccacatg 240  
DB 181 tggatgtgatagaacagtagtaagatcatctgcgcacgcatcactcctccacatg 240  
QY 241 gaacttgcgcaaatcttgcttctgcgctctacagttctctccacaaattttgttg 300  
DB 241 gaacttgcgcaaatcttgcttctgcgctctacagttctctccacaaattttgttg 300  
QY 241 gaacttgcgcaaatcttgcttctgcgctctacagttctctccacaaattttgttg 300  
DB 241 gaacttgcgcaaatcttgcttctgcgctctacagttctctccacaaattttgttg 300  
QY 301 acaaactaaatcttcaatctcatctctgcacaaagtgggttgaagaagatctcgc 360  
DB 301 acaaactaaatcttcaatctcatctctgcacaaagtgggttgaagaagatctcgc 360  
QY 301 acaaactaaatcttcaatctcatctctgcacaaagtgggttgaagaagatctcgc 360  
DB 301 acaaactaaatcttcaatctcatctctgcacaaagtgggttgaagaagatctcgc 360  
QY 361 aaacactaatcttctatctgttgcagtggttaacatcaacaaatcacaacatgttaa 420  
DB 361 aaacactaatcttctatctgttgcagtggttaacatcaacaaatcacaacatgttaa 420  
QY 421 aaaaaattcaacatttggatagatgtgtcacaatgtaagtgaactcttacaattt 480  
DB 421 aaaaaattcaacatttggatagatgtgtcacaatgtaagtgaactcttacaattt 480  
QY 481 taaacttacttaataatacggatctagcaatcttacttgcattaaaaattagaaga 540  
DB 481 taaacttacttaataatacggatctagcaatcttacttgcattaaaaattagaaga 540  
QY 541 gtcttctacttaaaaaaagaggaacacagagaggttgaacagagaaacggg 600  
DB 541 gtcttctacttaaaaaaagaggaacacagagaggttgaacagagaaacggg 600  
QY 601 tgcatagccttaagagcttcaactacttaaccccaaaccaagaagatgtcaacttcaa 660  
DB 601 tgcatagccttaagagcttcaactacttaaccccaaaccaagaagatgtcaacttcaa 660  
QY 661 caatctctctcccccgaacccgttttttgaacggtcagttcggagcagacacgt 720  
DB 661 caatctctctcccccgaacccgttttttgaacggtcagttcggagcagacacgt 720  
QY 721 taaggagcagctatattctcgtcttctcctctacacacatgcaiyccataataaag 780  
DB 721 taaggagcagctatattctcgtcttctcctctacacacatgcaiyccataataaag 780  
QY 781 ccggttgagatctttaaataataataataatacaagaagaaagctatttattcata 840  
DB 781 ccggttgagatctttaaataataataataataatacaagaagaaagctatttattcata 840

DB 781 ccggttgagatctttaaataataataataataatacaagaagaaagctatttattcata 840  
QY 841 agaaagaaagagagaaacacacacactaatcatagttcttcggaggttctt 900  
DB 841 agaaagaaagagagaaacacacacactaatcatagttcttcggaggttctt 900  
QY 901 gttgcgcttaataaaaagctcttcttattacttacttcaatgtagatttcccaaaa 960  
DB 901 gttgcgcttaataaaaagctcttcttattacttacttcaatgtagatttcccaaaa 960  
QY 961 gctctatttttttgaataaaaaaaagttcaatcatttcaacttctgtttacagt 1020  
DB 961 gctctatttttttgaataaaaaaaagttcaatcatttcaacttctgtttacagt 1020  
QY 1021 gttgtgtgagagagaggtgtgtgttgaagaaagacagacagacacgacgag 1080  
DB 1021 gttgtgtgagagagaggtgtgtgttgaagaaagacagacagacacgacgag 1080  
QY 1081 aattagatttttatttatttcttactcttctgttcttgaatgtaagtggttttaa 1140  
DB 1081 aattagatttttatttatttcttactcttctgttcttgaatgtaagtggttttaa 1140  
QY 1141 aagggtatcgaaaaaatgagtgagttgtgtgaggtgtctcgttaagtgtaag 1200  
DB 1141 aagggtatcgaaaaaatgagtgagttgtgtgaggtgtctcgttaagtgtaag 1200  
QY 1201 tgggtatttcggaaggttgaaggttctcgcagatcgaagacaaacaaagattcgaaa 1260  
DB 1201 tgggtatttcggaaggttgaaggttctcgcagatcgaagacaaacaaagattcgaaa 1260  
QY 1261 tttagcattgttttgaatgagtcgagttatgtgtgttcttacttacttactgat 1320  
DB 1261 tttagcattgttttgaatgagtcgagttatgtgtgttcttacttacttactgat 1320  
QY 1321 cttaactccgaatcattcaatgtgtgtgttcttctgaatattggaaggtgtgttact 1380  
DB 1321 cttaactccgaatcattcaatgtgtgtgttcttctgaatattggaaggtgtgttact 1380  
QY 1381 caattactcagcttactcttctcacttacttctcgaatcttcttcttatttgaagtg 1440  
DB 1381 caattactcagcttactcttctcacttacttctcgaatcttcttcttatttgaagtg 1440  
QY 1441 aatgcatacctttagtgcatttgaatcttgaatgaatgtgtgtcttcttactt 1500  
DB 1441 aatgcatacctttagtgcatttgaatcttgaatgaatgtgtgtcttcttactt 1500  
QY 1501 gtaagattttagtgcatttgaatcttgaatgaatgtgtgtcttcttactt 1560  
DB 1501 gtaagattttagtgcatttgaatcttgaatgaatgtgtgtcttcttactt 1560  
QY 1561 ttgtatattgaaagcagtttcttcttgaatgaatgtgtgtcttcttactt 1620  
DB 1561 ttgtatattgaaagcagtttcttcttgaatgaatgtgtgtcttcttactt 1620  
QY 1621 ttgtatattgaaagcagtttcttcttgaatgaatgtgtgtcttcttactt 1680  
DB 1621 ttgtatattgaaagcagtttcttcttgaatgaatgtgtgtcttcttactt 1680  
QY 1681 attggttatttattttagccttcttcttcaagagagccttgaatgaatgtgtcttctt 1740  
DB 1681 attggttatttattttagccttcttcttcaagagagccttgaatgaatgtgtcttctt 1740  
QY 1741 attgacttttaccacaaattgtagttaaattgttgaattgttcttcttgaatgtat 1800  
DB 1741 attgacttttaccacaaattgtagttaaattgttgaattgttcttcttgaatgtat 1800  
QY 1801 gcttgcatacttggaggttacttacttgaatgttgaatgaatgttcttgaatgtat 1860  
DB 1801 gcttgcatacttggaggttacttacttgaatgttgaatgaatgttcttgaatgtat 1860  
QY 1861 cctagcgttagtgaatccttcttgaatgttgaatgttgaatgttcttgaatgtat 1920  
DB 1861 cctagcgttagtgaatccttcttgaatgttgaatgttgaatgttcttgaatgtat 1920







DB 4081 A 4081

RESULT 2  
V06591 standard; cDNA to mRNA; 2089 BP.

AC V06591.03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
OS Arabidopsis thaliana.  
FT Key location/Qualifiers  
FT CDS 195..2072  
FT /tag= a  
FT /product= SERK protein

PN W09743427-A1.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI; 98-086529/08.  
DR P-PSDB; M47023.  
PT Production of apomictic seeds - useful in plant breeding

PS Claim 27, Pages 91-95; 123pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;

Query Match 13.2%; Score 539.2; DB 1; Length 2089;  
Best Local Similarity 85.1%; Pred. No. 3.5e-112;  
Matches 664; Conservative 0; Mismatches 3; Indels 113; Gaps 1;

QY 3302 cagccgcaagaagaatcagaagtcctcgggcaagcctaagaagttcttcttgccggagc 3361  
DB 1018 CTGCCGAAGAAGATCCAGAGTTCACTGGGACACCTCAAGAGGTTTCTTGGCGGAGC 1077  
QY 3362 tacaaatgctgagtgatggttaagtaacaagaacatttggcgagagtggttggga 3421  
DB 1078 TACAATGGCGAGTGTGTTAGTAACAAGACATTGGCGCAGAGTGGTTTGGGA 1137  
QY 3422 aagctacaaggagagccttgccagacagaaactctgttgctgctcaagaagactgaaggaag 3481  
DB 1138 AAGTCTACAAAGGAGACCTTGGCAGACGGAACCTGTGCTCAAGAGACTGAAGGAAG 1197  
QY 3482 agcgaaactcagtgagagagctccagtttcaagaagaagtagaagtagaagtagagag 3541  
DB 1198 AGCGAAGCTCAGGTGAGAGCTCCAGATTCAACAGAAAGTAGATGATGGCAG 1257  
QY 3542 ttcaatgaaacctgtgtgagattacgaggttctgtatgacacgagccgagagattgcttg 3601  
DB 1258 TTCATCGAAACCTGTGAGATTACGAGGTTTCTGTATGACACGACGAGAGATTGCTTG 1317

QY 3602 tgatccctacatgagcgaatggaagtgtgtctgtctcagagtgtaaaactaaacaa 3661  
DB 1318 TGTATCTTACATGGCCATGGAAGTGTGCTTGCTGCTCAG----- 1361  
QY 3662 ttaaacatctgtgtctctctcaactacttgaagcgtgttttcaatgttccct 3721  
DB 1361 ----- 1361  
QY 3722 ttatggttcataatgttgtttacaactaagacagagagccacgctcacaaactcc 3781  
DB 1361 ----- -AGAGGCCACCGCTCACAACTCC 1384  
QY 3782 gcttgatggccaacgagagagaatcgcgctagcgctcagctcaggtgttcttacct 3841  
DB 1385 GCTTGATGGCCAAACCGGAGAGATGCCCTTAGGCTCAGCTGAGTTGTCTTACT 1444  
QY 3842 acatgatactgcatcgcgaagatcatcaccgtgacgttaaaagcaacaactcctc 3901  
DB 1445 ACATGATCAGCTGCATCCGAGATCATTCACCGAGAGTAAGCAACAAACATCCTCT 1504  
QY 3902 agacgaagaatcgaagcggtgttgagagattcgggttgcaagtgtaataagctataa 3961  
DB 1505 AGACGAAGATTGGAAGCGGTTGGAGATTGGGTTGGCAAAAGCTTATGGAATAA 1564  
QY 3962 agaacatcagtgacaagaacgagtcgctggcaccatcggtacatcgctccaatatct 4021  
DB 1565 AGACATCTACGTCGACAAACAGCAGTCCGTGGCACCATCGGTCACTCCCTCCAGAAATCT 1624  
QY 4022 ctaacacggaataatctcaagagaanaacgacgcttctggatacgaatcatgctctaga 4081  
DB 1625 CTCACCGGAAATCTTGAGAGAAACCGACGATTTTCGATGCGAATCATGCTTCTAGA 1684  
RESULT 3  
V06570  
ID V06570 standard; DNA; 6695 BP.  
AC V06570.03-AUG-1998 (first entry)  
DE Daucus carota SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
OS Daucus carota.  
FT Key location/Qualifiers  
FT CDS 3696..6620  
FT /tag= a  
FT /note= "contains introns"  
FT Intron 3731..3802  
FT /tag= b  
FT /number= 1  
FT Intron 3851..3979  
FT /tag= c  
FT /number= 2  
FT Intron 4124..4211  
FT /tag= d  
FT /number= 3  
FT Intron 4284..4357  
FT /tag= e  
FT /number= 4  
FT Intron 4430..4528  
FT /tag= f  
FT /number= 5  
FT Intron 4642..4757  
FT /tag= g  
FT /number= 6  
FT Intron 4890..4967  
FT /tag= h  
FT /number= 7  
FT Intron 5295..5803  
FT /tag= i  
FT /number= 8  
FT Intron 6197..6339  
FT /tag= j  
FT /number= 9



PN MO9743427-21.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS.) NOVARTIS AG.  
PI De Vries SC, Hecht YFG, Schmidt EDL, Van Holst GJ;  
DR WPL; 98-086929/08.  
P-PSDB; W47013.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 21; Pages 40-46; 123PP; English.  
CC The sequence is that encoding SPKX, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and storage quality and maturity.  
SQ Sequence 6695 BP; 1844 A;T;C;G; 1243 G; 2422 T;

Query Match	9.5%;	Score 387.4;	DB 1;	Length 6695;
Best local Similarity	57.8%;	Pred. No. 7.3e-78;		
Matches 903; Conservative	0;	Mismatches 506;	Indels 152;	Gaps 6.

QY	2197	ttaaaggaggtttacagctaaacaactaactgcccagatctcttaatactttgaaatct	2256
Db	3974	TGACGGGAGGCTTTACAGCAATTAACATTAAGGTGACCAAATCTCTATGTATCTTGGAAATCT	4033
QY	2257	gacaacttagtgatttgattcttacttaacaacagcttctccggtccatctccgaatc	2316
Db	4034	GACAAATTTGGTGTGAGCTTTGGACCTTATACATGAATACTCTCTGTGACCTATACCGGAACT	4093
QY	2317	attgggaaggttccaagctgagattctctgtagatatacatgctttaccggctcagt	2376
Db	4094	ATTGGGAAGGTTTCANAGGCTTAAGATCTTGTAAGCTCAAAAT-----C	4138
QY	2377	tacagctcttctttaaactcttaggttttcttccaattttgaccttcttgtaaattta	2436
Db	4139	TTCAGTACTTTTAACTTAATGCAATTTGATTTATCTTTCAAGTGATGATTAATACCA	4198
QY	2437	catgcaagaataagccggcttacaacaacagtgctcactgtgtcattctcatctgctga	2496
Db	4199	AATTACTGGATAGGCGTCTAACAACAACAGCTCTCTGTGTCCAATTCGAATGTCACGA	4258
QY	2497	coaatattacaccccttcaagtgctgtagctctccatcaacttaactttaatgctact	2556
Db	4259	CTAATATTACAACGCTTCAAGTCTGAAATATTCGACCTTCCAGATAGTT-----	4312
QY	2557	tcattctccctcagtgatttgttgtagtaagcaacttaaccttgaatggaatgaacacag	2616
Db	4312	-----TTGTGTGTGGATGTTTAAATTTAATACATAATTAATGTTTCAATG	4357
QY	2617	agatctatcaataaacaagacctctgtgttcaagttccctgcaaatggtctcctcaactt	2676
Db	4358	GGATTTATCAACAACATGGCGCTATACAGAACCAATACCGGAATATGGGTCAATTTCTTGT	4417
QY	2677	cacaccctcagtgctctatgattatctcctctcaagtatttcaagtgtgtgtgcagtgct	2736
Db	4418	TTACACTTCACAGGTTTAATGCTACTAATATCTTTAATATTAATGATGTTCTTACTTACATGC	4477
QY	2737	tgaactttcttgaactttcaattc-----cttgtgaagtttgctaa	2780

Db	4478	GAAAGCTATGATTAATTAATTTTTTTTTCCTTCATTAATTAATCACTTTCGACGTTTTGGCAA	4557
Qy	2781	taactttagaactatgtagacctgttacaagtacaccatgtctcgtgatcccccgtttc	2840
Db	4538	TAATTTGAATTAATGATGACCTGTAACCTGGAGGCCCTGCCCTGATGATCCCCCAATTTTC	4597
Qy	2841	tctccacaccccttttattcaacctcccagtttccaccccgtagagcctctcttt	2900
Db	4598	TCCACACACCTCCGTTACTATCCACACATCAACAGTACAGCCTCCAGGTGATTAATTTTAT	4657
Qy	2901	tagtttaactatagagaacagaagaatgaaatccttgctctcctgtaatcccttttc	2960
Db	4658	ATTATTTCCCTATTAATTAATTTTATGACTGTAATAAATATGCTGTTAATTTACCAATGTTGGA	4717
Qy	2961	ataatacctatcttgccaataaggaataacaaatgataattgataatcagttggt	3020
Db	4718	ATTAAGTATTTTCCTCTTTT-----CTCTTCCTATTATTAATGAGACGAA	4763
Qy	3021	atggtataacatcgaggaatgctcgtgtgaggtgtcgtcaggtgtcgttgcttgcttggtg	3080
Db	4764	ATGCTCCACACAGCGACATTAATGCTGGGGAGTACGCTCTGCTGCTTACTGTTGCTG	4833
Qy	3081	ctctcgtcaatagcccttgccttgcttgggtggcagaagaagaagccactagataattctcagat	3140
Db	4834	CACCTGCATATGGCATTTGTCATGTGGGGGAGAGAAACCGGAGAACATTTCTTTGATG	4883
Qy	3141	tgccgtggtgattatataatccgattagttcttcgtcttagcagcaatttgtttgc	3200
Db	4884	TGCCGGTATCTCTGTAATTAATGATA-----	4910
Qy	3201	gaaagatttggaaacaactgttaatgaatacatataaagtcattgttttaagta	3260
Db	4910	-----TCTATGAAGCCT	4923
Qy	3261	caaatcttcttgatgaataactcgatgtgcgaatctctatgacgcggaagaatccga	3320
Db	4924	TACTGTCGTGGACTTTGTTTACGTCACTTATGTTAACTTACGTGAAGAGACCCGA	4983
Qy	3321	agttcatctggagacgtcacaagagtttcttgcggagctacaaagtgcgagtgatg	3380
Db	4984	AGTGACCTTGCTCACTGAAGAGGTTTCTGCGAGAAATGCAAGTGGCAACGAGATAC	5043
Qy	3381	gtttagttaacaagaacattttggcgaggttgggttgggaaagtctacaagagacgtt	3440
Db	5044	TTTATGATC-----CATCCTTGGAAAGGTGATTTGGATTAAGTGTATTAAGGAGCGCT	5097
Qy	3441	ggcagaaggaactcttgctgtcgaagagacgaagaagagcgaaatccaggtgtaga	3500
Db	5098	TGCTATATGCTCACTTGTAGCAGTTAAABGCTTTAAAGAAAGAACGACACCGAGTGGGA	5157
Qy	3501	gctccagtttcaaacagaagttagatgataagtaatgacgtatgcacatcgaacctgttag	3560
Db	5158	GCTGACAGTTTAAACACAGAGTGAAGTAAGATATACATGCGTGCATCGAAATCTTCTCG	5217
Qy	3561	acttagaggttctcgtagaacagcagcagagagattgctgtgtatccttaatagtgcaa	3620
Db	5218	TCTAGCTGTTTCTGATGACACTACCGAGCGGCTTTGTTGTTATCCATTAATGGCTAA	5277
Qy	3621	tggaaigtgtccttcgtgtctcagaggtaaanaactaaacaatnaacatctgtgtctc	3680
Db	5278	TGGAAGTGTGCGATCAGTTTAAAGGTATCTCAGTTACATTAACATTAATTCGACAGA	5337
Qy	3681	tctcaattactttagcaggaagtgtttttcaagtttctccttaatgagtgtaattgtt	3740
Db	5338	GTTTGTTGATTAATAAATGAATATTAACCTCCTTACACTATGTTAAAGGTGTTAATTTCT	5397
Qy	3741	g 3741	
Db	5398	G 5398	

## RESULT 4



V06571  
ID V06571 standard; cDNA; 1814 BP.  
AC V06571;  
DE 03-AUG-1998 (first entry)  
DE Daucus carota SERK gene.  
KW Receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KW plant breeding; ss.  
OS Daucus carota.  
FH Key Location/Qualifiers  
FT CDS 94..1755  
FT /tag= a  
FT /product= SERK protein  
PN W09743427-A1.  
PD 20-NOV-1997.  
PE 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS) NOVARRIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR MPI: 98-086529/08.  
DR P-PSDB; W47013.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 21; Pages 47-51; 123p; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilize a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

Query Match 6.6%; Score 270.8; DB 1; Length 1814;  
Best Local Similarity 64.6%; Pred. No. 7.6e-52;  
Matches 304; Conservative 0; Mismatches 157; Indels 119; Gaps 2;

QY 3302 cagccgaagaagatccagaagatccatctggaagcctcaagaggtttcttgcgggagc 3361  
DB 707 CAGCTGAAGAGGAGACCGAGAGTGCACCTGTGTAAGAGAGGTTTCTCTCGAGAAAT 766  
QY 3362 tacaagtgaggagtgatgaggtttagtaacaagaacatttggcagaaggtgggtttggga 3421  
DB 767 TGCAGAGTCGCAACGATACCTTTAGT-----ACCATACCTGGAAAGGAGTGGTA 820  
QY 3422 aagctcaagaagcgtctggaagcgaagcgaagcctgtgtctgtcaagaagcctcaaggaag 3481  
DB 821 AGGTGTAAGAGGACGCTTGCATGCTGACCTGTGTAAGAGCTTAAAGAGCTTAAAGAG 880  
QY 3482 agcgaactccagtgagagctccagttccaacaagaagtagagatagatagtcagag 3541  
DB 881 AACGAACACACAGTGTGATGCTGCAAGTTCMAACAGAGGTGAATGATTAAGCATGGCTG 940  
QY 3542 ttcatgaaactgttgagattagaggtttctgtatgaacacgagcgaagagattgttg 3601  
DB 941 TGCATCGAAATCTTCTGCGTCTAGCGGTGTTTCTGCATGACACCAACAGAGCGGCTTTTG 1000  
QY 3602 tgtatcttaccatgagcgaatggaagtggtgtgtctgtctcgaagtgtaaaactaaacaa 3661  
DB 1001 TATATCCATACATGCTAATGAGAGTGTGCGTGTGTTAAG----- 1044  
QY 3662 ttaaacatctgtgtctctctcaactttagacgtgaagtggtttttcatgtttctct 3721

DB 1044 ----- 1044  
QY 3722 ttatggttataatgtgtttacactaataacagagagccacgtccaaactcc 3781  
DB 1044 -----AGAGCGTCAGCCATCAGAACCTCC 1067  
QY 3782 gcttgatggccaacgcggaagagaatcgctaggtcgaagtcgaggtgttctact 3841  
DB 1068 CTTGATGGCCAACTGGAAGAGGATTCGATGAGTGTGCTAGGGGGCTTTATT 1127  
QY 3842 acatgacactgcgacccaagatcattccacgtgaagcgttaaaagcgaacacccctct 3901  
DB 1128 GCATGACCATTTGATTCACAGATATCCATCGATGTAAGAGCGCAATATATAT 1187  
QY 3902 agacgaagaatccgaagcgtgtgtgagattccggtgtggcgaagcgaatgactata 3961  
DB 1188 GGACGAAGAATTTGAGGCTGTGTGATGATTTGGGTGCTAGGCTCATGTATACAA 1247  
QY 3962 agacactcagtgacacaacagcagtcggtgcacatcggtcacatcgctccagaatct 4021  
DB 1248 GGAATCCCATTTACAACTGCTGAAGGGTACCTTGGCTATAGCTCCGAGTACCT 1307  
QY 4022 ctcaacggaatcttcagagaagaacacgagcgtttcgaatacgaatcgtctaga 4081  
DB 1308 CTCGACTGGAAGTCATCAGAGAGACCGATGCTTGTGTTAGGATATGCTTAGA 1367

RESULT 5  
T62124  
ID T62124 standard; cDNA to mRNA; 3176 BP.  
AC T62124;  
DE 10-JUN-1997 (first entry)  
DE Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.  
KW Plant; morphogenesis; regulation; short; stem; alteration;  
KW inflorescence; extraneous; gene; expression; transformation;  
KW increase; control; form; length; ds.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 51..2981  
FT /tag= a  
FT /note= "plant morphogenesis regulatory protein"  
PN J09056382-A.  
PD 04-MAR-1997.  
PE 24-AUG-1995; J61487.  
PR 24-AUG-1995; JP-216187.  
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.  
PA (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.  
DR MPI: 97-206629/19.  
DR P-PSDB; W13408.  
PT DNA encoding plant morphogenesis regulatory protein - useful to  
PT yield plants with short stems or altered inflorescence  
PS Claim 1; Pages 6-10; 17pp; Japanese.  
CC The present sequence encodes an Arabidopsis thaliana plant  
CC morphogenesis regulatory protein (MRP), which can be used to yield  
CC a plant with, e.g. short stems or altered inflorescence. The MRP  
CC acts on a plant at a specific site for a specific period, and can  
CC therefore be used to regulate extraneous gene expression in a  
CC plant. The MRP's cDNA or genomic DNA can be used to transform a  
CC plant to increase its MRP expression, and therefore control the  
CC form (particularly stem length) of the plant.  
SQ Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match 2.6%; Score 108; DB 1; Length 3176;  
Best Local Similarity 55.6%; Pred. No. 3.8e-15;  
Matches 207; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 3710 tcatgttcccttattggttcatatattgtgtttacactaataacagagagagccacc 3769  
DB 2201 TCTTCTGTTCTATGACTATTGGAATGATGACCTCTGGGATCTTCTTATATGAGCCCTTAC 2260  
QY 3770 gtcaaacctccgttgattggccaacgcggaagaagaatcgcttaggtcgaagtcgag 3829



```
Db 2261 GAAGAAAAAGACTCTGATTTGGACACACGGCTTAAAGATATGTTGGTGCACACAAAGG 2320
QY 3830 ttgtcttaccatcatgacatgcagtcgcgaagaatcaccgctgacgtaaaagcagc 3889
Db 2321 TTTAGCTTATCTACACCACTGACTGTGTCAGATCATCTACAGACAGCGTAAGTCTCTC 2380
QY 3890 aaacatccctcctagacgaagaattcgaaacggttcttggaagtttcgggttgccaaagct 3949
Db 2381 CAACATTCCTTTGGACAAAGACTTAGAGGCTGTTGACAGATTTTGGAAATAGCGAAAG 2440
QY 3950 aatgagactataaagacacacgacgacgacgacgacgacgacgacgacgacgacgacgac 4009
Db 2441 CTGTGTGTGTGTAAGTCACTACATCTTCACTTACGTGATGGCGACGATAGCTTACATAGA 2500
QY 4010 tccagaatctctcacaaccggaanaatctcagagaanaaacgacgcttttcgatacggaat 4069
Db 2501 CCCCAGATATGCTTCGACACTTCAACGGCTCAGTGAAGAAATCCGATGCTACAGTTATGAAT 2560
QY 4070 catgctctaga 4081
Db 2561 AGTCCCTTTTGA 2572
```

## RESULT 6

```
ID T62125 standard; DNA: 9295 BP.
AC T62125;
DT 10-JUN-1997 (first entry)
DE Arabidopsis thaliana plant morphogenesis regulatory protein DNA.
KW Plant; morphogenesis; regulation; short; stem; alteration;
   Inflorescence; extraneous; gene; expression; transformation;
KW increase; control; form; length; ds.
OS Arabidopsis thaliana.
FH Key
FT exon location/Qualifiers
   1803..1881
FT intron /tag= a
   1882..2227
FT exon /tag= b
   2228..2366
FT intron /tag= c
   2367..2467
FT intron /tag= d
   2540..2643
FT exon /tag= e
   2468..2539
FT intron /tag= f
   2644..2715
FT exon /tag= g
   2716..2809
FT intron /tag= h
   2810..2878
FT exon /tag= i
   2879..2968
FT intron /tag= j
   2969..3040
FT exon /tag= k
   3041..3118
FT intron /tag= l
   3119..3190
FT exon /tag= m
   3191..3266
FT intron /tag= n
   3267..3338
FT exon /tag= o
   3339..3421
FT intron /tag= p
   3422..3493
FT exon /tag= q
   3494..3586
FT intron /tag= r
   3587..3655
FT exon /tag= s
```

```
FT intron 3656..3740
FT /tag= t
FT exon 3741..3812
FT /tag= u
FT intron 3813..3888
FT /tag= v
FT exon 3889..3960
FT /tag= w
FT intron 3961..4048
FT /tag= x
FT exon 4049..4120
FT /tag= y
FT intron 4121..4209
FT /tag= z
FT exon 4210..4281
FT /tag= aa
FT intron 4282..4349
FT /tag= ab
FT exon 4350..4421
FT /tag= ac
FT intron 4422..4508
FT /tag= ad
FT exon 4509..4580
FT /tag= ae
FT intron 4581..4706
FT /tag= af
FT exon 4707..4778
FT /tag= ag
FT intron 4779..4860
FT /tag= ah
FT exon 4861..4932
FT /tag= ai
FT intron 4933..5018
FT /tag= aj
FT exon 5019..5090
FT /tag= ak
FT intron 5091..5176
FT /tag= al
FT exon 5177..5248
FT intron /tag= am
   5249..5412
FT exon /tag= an
   5413..5481
FT intron /tag= ao
   5482..5576
FT exon /tag= ap
   5577..5648
FT intron /tag= aq
   5649..5726
FT exon /tag= ar
   5727..5800
FT intron /tag= as
   5801..5882
FT exon /tag= at
   5883..6011
FT intron /tag= au
   6096..6443
FT exon /tag= av
   6012..6095
FT intron /tag= aw
   6444..6519
FT exon /tag= ax
   6520..6512
FT intron /tag= ay
   6891..6974
FT exon /tag= az
   6975..7328
FT /tag= ba
FT J09056382-A.
FT PD 04-MAR-1997.
FT PF 24-AUG-1995; 216187.
FT PR 24-AUG-1995; JP-216187.
FT PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
```



PA (CHK-) ZH CHIRYU KANKYO SANGYO GIITSU KENKYU.  
 DR WPI; 97-20629/19.  
 CC DNA encoding plant morphogenesis regulatory protein - useful to  
 PT yield plants with short stems or altered inflorescence  
 PS Claim 6; Pages 12-15; 17pp; Japanese.  
 CC The present sequence encodes an Arabidopsis thaliana plant  
 CC morphogenesis regulatory protein (MRP), which can be used to yield  
 CC a plant with, e.g. short stems or altered inflorescence. The MRP  
 CC acts on a plant at a specific site for a specific period, and can  
 CC therefore be used to regulate extraneous gene expression in a  
 CC plant. The MRP's cDNA or genomic DNA can be used to transform a  
 CC plant to increase its MRP expression, and therefore control the  
 CC form (particularly stem length) of the plant.  
 SQ Sequence 9295 BP; 2873 A; 1581 C; 1602 G; 3239 T;

Query Match 2.6%; Score 106.2; DB 1; Length 9295;  
 Best Local Similarity 58.3%; Pred. No. 1.4e-14;  
 Matches 166; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 3763 ggcacccgtcaaacctcgcgtgattggcacaacgcggaagatcgctgctcag 3822  
 DB 6520 GCCCTACGAAGAAAGACCTTGATGGACACACGCGCTTAAGATAGCATATGATGACAG 6579  
 QY 3823 ctgcaggttgctctaccctacatgacatcgatcgacgaagatcattcaccgtagctaa 3882  
 DB 6580 CACAAAGTTTAGCTTATCTTACACCATGACTGTAGTCCAAAGATATTACAGAGACGTGA 6639  
 QY 3883 aagcagaagaacatcctcttagacgaagaattcgaagcgttgcttgagattcgggttgg 3942  
 DB 6640 AGTGTCCACATCTCTTGACAAAGACATTAGAGCGTCTTGACAGATTGGAATAG 6699  
 QY 3943 caaagcgaatggaataaagacacatcagctgacaaacagcagcctggtgacatcgtgc 4002  
 DB 6700 CGAAAGCTTGCTGTGTCAAGTCAACATCTTCACTTACGTGATGGACATGAGTT 6759  
 QY 4003 acatcgctccagaatctctcacaacgcgaaatcttcagagaagaacgcagcttcggat 4062  
 DB 6760 ACAATAGACCCGAGTATGCTCGCACTTCACGGCTCAGCTCATGAGAAATCCGATGTACAGTT 6819  
 QY 4063 acggaatcgtctctaga 4081  
 DB 6820 ATGGAATAGTCTTCTTGA 6838

RESULT 7  
 X07356  
 ID X07356 standard; DNA; 4104 BP.  
 AC X07356;  
 DE Arabidopsis steroid receptor Bin1 DNA.  
 KW BIN1; steroid receptor; receptor kinase; transgenic plant;  
 KW brassinosteroid; disease resistance; crop protection;  
 OS Arabidopsis thaliana.  
 FH Key Location/Qualifiers  
 FT CDS 97..3687  
 FT 3687  
 FT /tag= a  
 FN W09859039-A1.  
 PD 30-DEC-1998.  
 PE 24-JUN-1998; U13100.  
 PR 24-JUN-1997; US-881706.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PI Chory T, Li J;  
 DR WPI; 99-081275/07.  
 DR P-PSD; W97819.  
 PT New receptor kinase Bin1 involved in brassinolide signalling -  
 PT useful for promoting increased yield and disease resistance in  
 PT plants and for modulating oocyte maturation  
 PS Claim 8; Page 49-52; 72pp; English.  
 CC This DNA sequence codes for a novel plant steroid receptor kinase,  
 CC designated Bin1 (see W97819), which is involved in the pathway for  
 CC the synthesis of the plant steroid hormone, brassinolide. 18 New

CC Arabidopsis dwarf mutants were identified that lacked the ability  
 CC to respond to brassinolide, and were named bin mutants. The bin1  
 CC mutations were used to map the gene to a small interval on  
 CC Arabidopsis chromosome 4. Bin1 was cloned by map-based cloning.  
 CC The Bin1 polynucleotide was identified within this interval by  
 CC sequencing the wild-type and mutant alleles of this nucleic acid.  
 CC Overexpression of Bin1 in transgenic plants provides plants  
 CC characterised as having enhanced disease resistance, increased  
 CC plant yield or vegetative biomass and increased seed yield.  
 CC Expression of Bin1 may also increase resistance to pesticides.  
 CC Inhibition of Bin1, e.g. using antisense oligonucleotides (AON), is  
 CC used to render plants male-sterile, and to reduce their stature or  
 CC yield, e.g. for creating dwarf varieties. Since Bin1 homologues  
 CC may be involved in regulation of the menstrual cycle and uterine  
 CC function, Bin1, antibodies and AON may be useful as contraceptives,  
 CC for improving success of in vitro fertilisation and to prevent  
 CC premature labour. Transgenic animals are also provided, and are  
 CC models for studying steroid-receptor interactions or can be used  
 CC to screen for therapeutic agents.  
 SQ Sequence 4104 BP; 1061 A; 870 C; 930 G; 1243 T;

Query Match 2.1%; Score 87.6; DB 1; Length 4104;  
 Best Local Similarity 57.9%; Pred. No. 1.6e-10;  
 Matches 175; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 3763 ctgtattggccaacgcggaagatcgctgctagcctcagctgagttgcttactcta 3842  
 DB 3028 CTTAAATTTGTCACACGCGGGAAGATTGCGATGAGATCAGTAAAGGCTTGCTTCTT 3087  
 QY 3843 catgatacctcgatcgaagaatcattcaccgttagctaaagaagaacaatcctcta 3902  
 DB 3088 CACCAACATGTGAGTCCCATATCATCATCAGAGAACATGAAATCCAGTAATGTTGCTT 3147  
 QY 3903 gacgaagaattcgaagcgttgcttgagattcgggttgagcaagcgaactaagactataa 3962  
 DB 3148 GATGGAATTTGGAAAGTCTGGGCTTTCAGATTTTGGCATGGGAGCTGATGAGTGCATG 3207  
 QY 3963 gacatcgcgga---caacagcagtcgctgagcacaatcggctcaccgtccagaatat 4019  
 DB 3208 GATACGACATTTAAGCGGTGATGATACCTAGCTGTACACGGGTTACGTTCTCCAGAGTAT 3267  
 QY 4020 ctctcaacgcgaaatctctcagagaagaacgcagcttcgcgatacggaaatcgtctcta 4079  
 DB 3268 TACCAAGTTTCAGGTGTTCAACAAGAGACGTTTATGTTACGGGTGCTTCTTCTC 3327  
 QY 4080 ga 4081  
 DB 3328 GA 3329

RESULT 8  
 X23531  
 ID X23531 standard; cDNA; 3045 BP.  
 AC X23531;  
 DE 17-JUN-1999 (first entry)  
 DE Maize Kaz1 gene DT4 cDNA fragment.  
 KW Kaz1; receptor kinase-like protein; multigene family; RRK; rice; DT4;  
 KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
 OS Zea mays.  
 PN W09909151-A2.  
 PD 25-FEB-1999.  
 PE 17-JUL-1998; U14841.  
 PR 13-AUG-1997; US-910386.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 DR WPI; 99-204431/17.  
 PT New RRK polynucleotides and nucleic acid constructs - used for  
 PT generating transgenic plants resistant to Xanthomonas  
 PS Claim 12; Page 58-59; 67pp; English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins











DR WPI, 99243728/20.  
 Pt New apoptosis-resistant virus-sensitive cell  
 Ps Example 3, Page 46-49; 51pp; English.  
 Cc The present invention describes an apoptosis-resistant virus-sensitive  
 Cc cell line into which an apoptosis resistance gene has been introduced.  
 Cc The recombinant viruses generated are capable of expressing apoptosis-  
 Cc associated genes. These can then be used in a variety of diseases for  
 Cc which the induction of apoptosis is gene transfer, or where the  
 Cc inhibition of harmful apoptosis is therapeutic. The recombinant viruses  
 Cc are useful as vectors for gene therapy which can be applied to cancer  
 Cc therapy for destroying cancer cells selectively, the treatment of  
 Cc autoimmune diseases and graft rejection reaction, and apoptosis induction  
 Cc therapy for inflammatory cells in inflammatory diseases. Prior arts have  
 Cc encountered the problem where if an adenovirus vector capable of  
 Cc expressing an apoptosis-associated gene is introduced into animal cells,  
 Cc the cells producing the virus will be destroyed because the period of  
 Cc time required to induce cell death by apoptosis is shorter than that  
 Cc required to replicate and produce the virus, resulting in failure to  
 Cc obtain a recombinant virus having the integrated apoptosis-associated  
 Cc gene. In this invention an apoptosis-resistant 293 cell line (having an  
 Cc apoptosis resistant gene introduced) is established and overcomes the  
 Cc problem. The present sequence represents the base sequence of the  
 Cc plasmid ptx-Bcl 2-1-hoc 35, which contains the human Bcl-2 gene, and  
 Cc is used in an example from the present invention.  
 Sg Sequence 7996 BP; 2463 A; 2015 C; 1829 G; 1689 T;

Query Match	2.0%	Score 82.2;	DB 1;	Length 7996;
Best Local Similarity	45.5%;	Pred. No. 3.4e-09;		
Matches 291; Conservative	0;	Mismatches 348;	Indels 0;	Gaps 0

DT 17-JUN-1999 (first entry)  
 DE Tomato Xa21 clone TRK2 DNA fragment.  
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice; TRK2;  
 KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
 OS Lycopersicon esculentum.  
 PN W09090151-A2.  
 PD 25-FEB-1999.  
 PF 17-JUL-1998: U14841.  
 PR 13-AUG-1997: US-910386.  
 RA (RSCC ) UNIV CALIFORNIA.  
 PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 DR WPI: 99-204431/17.  
 DR P-PSDB: W23533.  
 PT New RRR polynucleotides and nucleic acid constructs - used for  
 generating transgenic plants resistant to Xanthomonas  
 Claim 14: Page 61-62: 67pp: English.  
 PS This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the Oryza longistylis and Oryza sativa receptor  
 CC kinase-like protein (RRK) Xa21 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:42:10 ; Search time 274.21 Seconds  
(without alignments)  
1934.536 Million cell updates/sec

Title: US-09-180-798-20

Perfect score: 4081  
Sequence: 1 tctagaacatttgatcat.....tacggaatcatgcttctaga 4081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/5C.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/5D.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/6.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/PCUS.COMB.seq:\*  
7: /cgn2\_6/ptodata/2/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72.2	1.8	7218	1 US-08-232-463-14	Sequence 14, Appl
2	71.6	1.8	1554	4 US-08-587-680A-24	Sequence 24, Appl
3	64	1.6	5733	3 US-08-473-553A-1	Sequence 1, Appl
4	59.8	1.5	966	1 US-08-447-185-2	Sequence 2, Appl
5	59.8	1.5	2443	1 US-08-447-185-3	Sequence 3, Appl
6	59	1.4	2571	1 US-07-717-331F-9	Sequence 9, Appl
7	59	1.4	2833	1 US-07-717-331F-1	Sequence 1, Appl
8	55.4	1.4	2749	1 US-07-717-331F-4	Sequence 4, Appl
9	54.2	1.3	2749	2 US-08-265-628-1	Sequence 1, Appl
10	53.2	1.3	19124	4 US-08-487-826B-13	Sequence 13, Appl
11	47.8	1.2	240	1 US-08-628-417-6	Sequence 6, Appl
12	46.2	1.1	3921	4 US-08-567-375-3	Sequence 3, Appl
13	46.2	1.1	3921	4 US-08-587-680A-3	Sequence 3, Appl
14	46.2	1.1	5392	3 US-08-475-891A-3	Sequence 3, Appl
15	46.2	1.1	6256	3 US-08-475-891A-1	Sequence 1, Appl
16	46.2	1.1	6256	4 US-08-567-375-1	Sequence 1, Appl
17	46.2	1.1	6256	4 US-08-587-680A-1	Sequence 1, Appl
18	46.2	1.1	2447	4 US-09-014-969-14	Sequence 14, Appl
19	45.4	1.1	7218	1 US-08-232-463-14	Sequence 14, Appl
20	45.4	1.1	6243	4 US-09-056-075-1	Sequence 1, Appl
21	44.4	1.1	13473	6 PCT-US96-03916-1	Sequence 1, Appl
22	44.4	1.1	18912	6 PCT-US96-03916-59	Sequence 59, Appl
23	43.8	1.1	2621	4 US-08-553-619B-8	Sequence 8, Appl
24	43.2	1.1	1611	1 US-07-820-011A-3	Sequence 3, Appl
25	43.2	1.1	1611	6 PCT-US93-00445-3	Sequence 3, Appl
26	42.8	1.0	19124	4 US-08-487-826B-13	Sequence 13, Appl
27	41	1.0	5852	1 US-07-867-106-2	Sequence 2, Appl

28	40.6	1.0	2208	6 PCT-US95-08493-1	Sequence 1, Appl
29	40.6	1.0	2580	6 PCT-US95-08493-18	Sequence 18, Appl
30	40.6	1.0	2604	6 PCT-US95-08493-20	Sequence 20, Appl
31	40.6	1.0	2869	1 US-08-374-834-2	Sequence 2, Appl
32	40.6	1.0	2869	2 US-08-644-271-2	Sequence 2, Appl
33	40.4	1.0	2216	2 US-08-046-583-12	Sequence 12, Appl
34	40.4	1.0	2216	3 US-08-384-556A-3	Sequence 3, Appl
35	40.4	1.0	2216	3 US-08-331-355A-24	Sequence 24, Appl
36	40.4	1.0	2216	6 PCT-US94-12364-24	Sequence 3, Appl
37	40.4	1.0	2216	6 PCT-US95-07553-3	Sequence 3, Appl
38	40.2	1.0	642	2 US-08-764-100-13	Sequence 13, Appl
39	40.2	1.0	3000	2 US-08-764-100-9	Sequence 9, Appl
40	39.8	1.0	1920	1 US-08-186-222-1	Sequence 1, Appl
41	39.6	1.0	831	4 US-08-567-375-15	Sequence 15, Appl
42	39.4	1.0	3641	1 US-08-030-096-5	Sequence 5, Appl
43	39.4	1.0	5775	2 US-08-306-691B-15	Sequence 15, Appl
44	39.4	1.0	5775	6 PCT-US93-06251-29	Sequence 29, Appl
45	38.6	0.9	1058	1 US-08-238-163-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7216 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F1s  
; US-08-232-463-14



Query Match 1.8%; Score 72.2; DB 1; Length 7218;  
Best Local Similarity 7.3%; Pred. No. 1.8e-08;  
Matches 32; Conservative 237; Mismatches 170; Indels 0; Gaps 0;

QY 2462 caacagctctcctggtgcaatcctctagtcacgacacatactacaccttaagtgt 2521  
DB 1011 CCAATACGCTCAAGAAATTAATTCGAGCTTGCTGCAAGTGAAGGAGCTTGCGATATTT 1070  
QY 2522 gtgagctcctcattcaacttctatctgctcactcctccctcagttgagttgtg 2581  
DB 1071 YY 1130  
QY 2582 agttaatgcaacttaaccttgatgagcaacagagatctatcaataacacctct 2641  
DB 1131 YY 1190  
QY 2642 gttcagctcctcgaacgtgctcctcctcactcctcaccacctcagttctatgatta 2701  
DB 1191 YY 1250  
QY 2702 tccctcctcagttatctcagttgtgctcagttgtcgaactatcttgaaccttca 2761  
DB 1251 YY 1310  
QY 2762 ccttgagagtttgctcctaacttagacactatgtagcttgtagacagtcaccatg 2821  
DB 1311 YY 1370  
QY 2822 ctgagatcccccgtttctcctccacacaccttattcaacctccccagttcacccc 2881  
DB 1371 YY 1430  
QY 2882 cgaagtaagcctcctcttt 2900  
DB 1431 YYYGTACCAATTCCTCT 1449

RESULT 2  
US-08-587-680A-24

; Sequence 24, Application US/08587680A

; Patent No. 5977434

; GENERAL INFORMATION:

; APPLICANT: Ronald, Pamela C.

; APPLICANT: Wang, Guo-Liang

; APPLICANT: Song, Wen-Yuang

; APPLICANT: Szabo, Veronique

; TITLE OF INVENTION: Procedures and Materials for Conferring

; TITLE OF INVENTION: Disease Resistance in Plants

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/587,680A

; FILING DATE: 17-JAN-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/373,375

; FILING DATE: 17-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/475,891

; FILING DATE: 07-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/004,645

; FILING DATE: 29-SEP-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/567,375

; FILING DATE: 04-DEC-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Bastian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 023070-0589400S

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1554 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA (partial)

US-08-587-680A-24

Query Match 1.8%; Score 71.6; DB 4; Length 1554;  
Best Local Similarity 54.6%; Pred. No. 1.2e-08;  
Matches 165; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 3783 ctgattggccaagcggaagagatcgagctcagctcagcgagttgtcttacctta 3842  
DB 982 CTCGAATGGAAATAGCTAAATATGCAATGAGCTGCCAAGGCTTCTTATTTG 1041  
QY 3843 catgatactgagtcgaagacatcaccgctgacgttaaaagcagcaacatcctta 3902  
DB 1042 CACCAAGATGCTCCCTATGATATTCACCCGAGTCAAGTCCAAATATTTGTTG 1101  
QY 3903 gacgaagatcgagcggtgtgttgagatttgggttgcaag---ctaagacct 3959  
DB 1102 AACTGTGAAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1161  
QY 3960 aaagacacatcgagcgaagacagagtcgagcaccatcgatcacatcgctccagaat 4019  
DB 1162 GGTACTCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1221  
QY 4020 ctctcaacggaataatcttcaagaaacagagcttctgagtaaggaatctgtctta 4079  
DB 1222 GCATACAGCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1281  
QY 4080 ga 4081  
DB 1282 GA 1283

RESULT 3

US-08-473-553A-1

; Sequence 1, Application US/08473553A

; Patent No. 5859338

; GENERAL INFORMATION:

; APPLICANT: Meyerowitz, Elliot M.

; APPLICANT: Clark, Steven E.

; APPLICANT: Williams, Robert W.

; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,

; TITLE OF INVENTION: Transformed Plants, and Proteins

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30



```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SILVA, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
; US-08-473-553A-1

Query Match          1.6%; Score 64; DB 3; Length 5733;
Best Local Similarity 53.6%; Pred. No. 1.8e-06;
Matches 133; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 3712 attgttccttaaggttcataatgtgtgtacactaatgacacagagagccaccgt 3771
DB 4733 atctctcttttatgatacatgcttaattggaagccttgagagcttttgatgatgata 4792
QY 3772 cacaaactcgcctgattgagcacaagcggaaagagaatcgcgctagctcagctcgaagtt 3831
DB 4793 aaggtggtcatcttcattgagagagacatgagatgagccgtggaacctgcacaaaggcct 4852
QY 3832 tctctactacatgatactcgcgcgcgaagatcattcacacgctgagctaaagcagcaa 3891
DB 4853 tctgttatcttaccatgattgttccaccattgatttgcataagacattttaagtcacata 4912
QY 3892 aatccctcttagaagaagaattcgaagcggttgttggagatttcgggttggcagaagctaa 3951
DB 4913 acattctttggacactgattttgaagcccatgctgctgattttggcctgtctaagtct 4972
QY 3952 tggactcat 3959
DB 4973 tagttgat 4980

RESULT 4
US-08-447-185-2
; Sequence 2, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATHOGE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,078
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman M., Michael L.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1000
; TELEFAX: (716) -263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-447-185-2
```

```

Query Match          1.5%; Score 59.8; DB 1; Length 966;
Best Local Similarity 53.4%; Pred. No. 8.1e-06;
Matches 149; Conservative 0; Mismatches 127; Indels 3; Gaps 1;

QY 3347 ttctcttgaggagacacaaagtcgagtgatgagtttaagtaagaagacatttgagca 3406
DB 86 ttcttttaagtagatttggggaagacaaactaaatatttgatcacaagttttaattggac 145
QY 3407 gaagtgagtttggaaagctacaagggagccttgagcagcgaactctgttctgca 3466
DB 146 atgctgtcttttgggaagcttttaagggctgttccgcatgagacacaagctgcccctga 205
QY 3467 agagacttgaaggaagagcgaactcgaagtggagagctccagtttcaacagaagtagaga 3526
DB 206 aaagcggtacacctgactcctcacaaagatattgaag--agttcgaaacagaaattggaga 262
QY 3527 tgaatgattgcaagtctcgaacacctgttgagattcagaggttctgtagcacccga 3586
DB 263 ctcttcattttggacagacatccgcatctggttctattatagagattcgtgataagaa 322
QY 3587 ccgagagattgcttgtatccttaccatgagccaatggaa 3625
DB 323 atgagatgatttcaatttaataatcatgacagaaatggga 361

RESULT 5
US-08-447-185-3
; Sequence 3, Application US/08447185
; Patent No. 5648599
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Martin, Gregory B.
; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
; TITLE OF INVENTION: TO PLANTS BY RESPONDING TO AN AVIRULENCE GENE IN PLANT PATH
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael L. Goldman
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,185
```



```

:      REGISTRATION NUMBER: 26,824
:      TELECOMMUNICATION INFORMATION:
:      TELEPHONE: (203)268-1951
:      TELEFAX: (203)268-1951
:      INFORMATION FOR SEQ ID NO: 9:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 2571 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: DNA
:      US-07-717-331F-9

```

Db	1859	CTAAGCTAAATTTGGAAATAGAGATTCCAGACNTTACCAATAGGTTGCTCTGAGGCGTTTAT	1918
Qy	3838	acctcatgatcatcgtcgatccgaagatcattcacgcgtgacgtlaaagcagcaaacatcc	3897
Db	1919	ATCTTCATCAACAGACTCCACGGTTTATAGATTAATCCACAGAGATTGAAAGTTAAGTAACTATT	1978
Qy	3898	tcttagcagaagaattcgaagcaggtgtgttggagatttcagggttggcaagctaagtgact	3957
Db	1979	TGCTTGGACAAAATATGATATCCCAAGATCTCGGATTTTGGGATTTGGCAGAGATATTGGAA	2038
Qy	3958	ataaagaacctcagctgacaacagca---gtccgtggcaccatcggtgacatcgctcaag	4014
Db	2039	GGGAGGAAACGGAACTTAACTAACATGAAGGTGGTGGGAACTATCCGCTTACATGTGCCGG	2098
Qy	4015	aatatcttcaaacccggaatatcttcagaagaaacccgagtttcttggagatcggaaatcagc	4074
Db	2099	AATACGCAATGTATGTGGATATTCTTCGGAAAAATCATGATGTTTTCAGTTTGGAGTCATAG	2158
Qy	4075	ttctaga 4081	
Db	2159	TTCTTGA 2165	

Sequence 1, Application US/07717331F  
Patent No. 5484905  
GENERAL INFORMATION:  
Applicant: June Nasrallah; Michael Nasrallah; and Joshua  
Applicant: Stein  
TITLE OF INVENTION: A Receptor Protein Kinase Gene  
TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Yahwak & Associates  
STREET: 25 Skytop Drive  
CITY: Trumbull  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06611  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/717,331F  
FILING DATE: June 19th 1991  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: George M. Yahwak  
REGISTRATION NUMBER: 26,824  
TELECOMMUNICATION INFORMATION:



TELEPHONE: (203)268-1951  
TELEFAX: (203)268-1951  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2833 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-07-717-331F-1

Query Match 1.4%; Score 59; DB 1; Length 2833;  
Best Local Similarity 51.8%; Pred. No. 2.2e-05;  
Matches 159; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 3778 ctcccttattgcccacgcgcgaagagacgcgcctagcctcagctcgagttgtctt 3837  
DB 1859 cttagcttaattggaattggaatgcgacatttaccattggtgttcgcagggcttttat 1918  
QY 3838 acctacatgatactgcgataccgaagatcatcaccgtgaagcgttaagcagcaacatcc 3897  
DB 1919 attcttattcgaacacgtcagcgtttgagatattccacagagattgamaagtaagtaacattt 1978  
QY 3898 tcttaagcagaagattcgaaagcgttgttgagatttcgaggttggaagcagtaagtact 3957  
DB 1979 tgccttgcacaaaattatgattcccaaaagattctcgattttggcattggccacgattattgaaa 2038  
QY 3958 ataagaacatcagtgacaacagca--gtccgtgacacatcggtcacatcgctccag 4014  
DB 2039 ggagacgaacggagacgaacacatgaaagagtgctggaacatagcgcctacgtcccgcg 2098  
QY 4015 aatatctcacaacgcgaatcttaagagaacacgcagcgtttcgagtaagcagatcgtc 4074  
DB 2099 aattaccatgattgagattctcgcaaaaatcagattgttttcagtttggaatcattag 2158  
QY 4075 ttctaga 4081  
DB 2159 ttcttga 2165

## RESULT 8

US-07-717-331F-4  
Sequence 4, Application US/07717331F  
Patent No. 5484905  
GENERAL INFORMATION:  
APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua  
TITLE OF INVENTION: A Receptor Protein Kinase Gene  
TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: yahwak & Associates  
STREET: 25 Skytop Drive  
CITY: Trumbull  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06611  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/717,331F  
FILING DATE: June 19th 1991  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: George M. Yahwak  
REGISTRATION NUMBER: 26,824  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203)268-1951  
TELEFAX: (203)268-1951

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2749 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-07-717-331F-4

Query Match 1.4%; Score 55.4; DB 1; Length 2749;  
Best Local Similarity 53.2%; Pred. No. 0.00017;  
Matches 141; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 3394 aacatttggcagagatgggttcttggaagatctacaagagacgttggcagcaggaact 3453  
DB 1612 AACCAAGCGCGGAAAAAGGGGTTTGCTGTGTTTCAAGGGAAGGTTAGTGAAGCGGCAA 1671  
QY 3454 ctgttgcttcaagaagactgaagaaagcgaactcaggtggagagctcagttcaa 3513  
DB 1672 GAATTTGCAATGGAAGAGACTATCGAAATGTCAGCTCAAGGTACCGATG--AGTTCAAG 1728  
QY 3514 acagaagtagaagatgataagatgacagttcatcgaacacctgttgagatgacagagttc 3573  
DB 1729 AACGAAGTTAGGCTATGCAAAAGCTTCAGCCACAAATATCTTCCGACTTCTGGCTGT 1788  
QY 3574 tgatgacacgcgcgaagagatgcttggtatccttacaatgccaatggaagtgttgt 3633  
DB 1789 tctgtttatgagggcgaagaaagatttatttttgcagatcttgagatttgaagatttgaaccttcgat 1848  
QY 3634 tctgtctcagaaggttaaaactaa 3658  
DB 1849 tctcatctctttgatgaacacgaa 1873

## RESULT 9

US-08-265-628-1  
Sequence 1, Application US/08265628  
Patent No. 5821094  
GENERAL INFORMATION:  
APPLICANT: Rothstein, Steven J.  
TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A  
TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 W. Madison St. Suite 3400  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/265,628  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/959,945  
FILING DATE:  
APPLICATION NUMBER: US 07/847,564  
FILING DATE: 03-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Poehoplen Ph.D., Donald J.  
REGISTRATION NUMBER: 32,167  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-707-8889  
TELEFAX: 312-707-9155  
INFORMATION FOR SEQ ID NO: 1:



SEQUENCE CHARACTERISTICS:  
LENGTH: 2749 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Brassica napus  
STRAIN: oleifera  
INDIVIDUAL ISOLATE: W1  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: S-locus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2574  
PUBLICATION INFORMATION:  
AUTHORS: GORING, DAPHNE  
AUTHORS: ROTHSTEIN, STEVEN J.  
TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A  
TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A  
TITLE: FUNCTIONAL SERINE/THREONINE KINASE  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749  
US-08-265-628-1

Query Match 1.3%; Score 54.2; DB 2; Length 2749;  
Best Local Similarity 50.8%; Pred. No. 0.0034;  
Matches 156; Conservative 0; Mismatches 148; Indels 3; Gaps -1;

QY 3778 ctccgcttgattggccaaagcggaagaatgcgctagcgcctcgaagttgctt 3837  
DB 1862 CTACGTTAAATGGAGACAGACATTTCACATTCACAAATGGTGTGCTCGAGACTTTAT 1921  
QY 3838 acctaatatcatcagctcgatccgaagatcatcaccgtgacgttaaaagcaaacatcc 3897  
DB 1922 ATTCTCATCAGACTCAGCGTTTGAAGTATCCACAGAGATATGAAGTAAGTAACTTT 1981  
QY 3898 tctttagaagaaatcgaagcggtgttgagagattcggtgttgcaagaactaaggact 3957  
DB 1982 TCGTTATATAAATATGACACCAAGATCTCGGATTTTGGGATGGCCAGAAATCTTTCNA 2041  
QY 3958 ataagaacctcagctgacacacagca---gtccgtgacacacgtgcacatcgctccag 4014  
DB 2042 GGAACAGAGCTGAAGCTAACACAGAGAGTGTGCGAACTTACGCTACATGTCTCCGG 2101  
QY 4015 aatactctcaaccggaataatcttcagagaaacccgacgttttcggatagcgaaatcgc 4074  
DB 2102 AGTACCAATGATGGGTATCTCGAATAAATCAGATGTTCAGTTTGGAGTCATTG 2161  
QY 4075 ttctaga 4081  
DB 2162 TTCTTGA 2168

RESULT 10

US-08-487-826B-13/c  
Sequence 13, Application US/08487826B

GENERAL INFORMATION:  
PATENT No. 5993827  
APPLICANT: Slim, Kim L.  
APPLICANT: Chitnis, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhaun  
APPLICANT: Wellens, Thomas E.  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487.826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ. ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 1.3%; Score 53.2; DB 4; Length 19124;  
Best Local Similarity 48.1%; Pred. No. 0.0016;  
Matches 180; Conservative 0; Mismatches 193; Indels 1; Gaps 1;

QY 1336 ttcacatgctgcttcttcctgaattggaaggttcggtgttactcaattactcaagcttt 1395  
DB 15784 TTTTCTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 15725  
QY 1396 actcgtttcccaattactcttcgatctcttttatttggaagtgatcgctctcttg 1455  
DB 15724 ATTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 15665  
QY 1456 tcttcgcatcttgatctgaatgaaatgtgtgttcttccttgatcttgaaattagtcgc 1515  
DB 15664 TTCTCTCTTTTGTATTATTTTATTTTATTCATTTTATTTTATTTTATTTTATTTTATTT 15606  
QY 1516 tagtacttgaaatacactgtttgtcttcttcgttcagatcaacttgatattgtaag 1575  
DB 15605 TAAATTTTATTTGAATATCTTTTCAATTTTATTTCTATCAAAATTTAATTTAATTTAATA 15546  
QY 1576 gcaatgtcttgggttgaagcgcgggttatattgatattcctaagatgatgttgatc 1635  
DB 15545 ATTTTATTAATTTTAAATAATTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 15486  
QY 1636 caaacatctctgaagactcattgttttggttttgaaagaaattgttaattatc 1695  
DB 15485 AAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15426  
QY 1696 agccctaatctca 1709  
DB 15425 CTTTATTAATTAATA 15412

RESULT 11

US-08-628-417-6/c  
Sequence 6, Application US/08628417

PATENT No. 5627054  
GENERAL INFORMATION:  
APPLICANT: GILLESPIE, DAVID  
TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC  
TITLE OF INVENTION: POLYMERASE CHAIN REACTION



```

Query Match          1.1%; Score 46.2; DB 4; Length 3921;
Best Local Similarity 49.4%; Pred. No. 0.041;
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2114 acaanaacaatccaagatttgacaagaagcaactgagatthacccttttgaatgaatc 2173
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1054 AAATTCACAACTTGAACCTGGGAAAAATACCTGGGGGGAGTTCTCTAATTCGTTT 1113
QY 2174 ttttaacaagttctctatttctctacaaggagctttacagtaaacacataactggccg 2233
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1114 TCCCACTTTCACCTTCGCTTAGTTTCCTTCGCACTTGAATTGAATGAATCAACAGAAAC 1173
QY 2234 attcctagtaatcttgygaactctgacaaacttgagttgagttcttacttaacagc 2293
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1174 ATTCCGAGAGATATTTGGCACTCTATTGGCTTACACATCTCTATCTGCAACAACAAAT 1233
QY 2294 ttcccggtctctatttcgcgaatcatctggaaagcttccaagctgaaattctgtgagta 2353
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1234 TTCAGAGGAGCTTTCCACATCATGTTGGGCAAGCTTAAATAACTTAAAGCAATTCATCGCC 1293
QY 2354 tac 2356
      |||
Db 1294 TAC 1296

RESULT 13
US-08-587-680A-3
; Sequence 3, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:

```



APPLICANT: Ronald, Pamela C.  
 APPLICANT: Wang, Guo-liang  
 APPLICANT: Song, Wen-Yuang  
 APPLICANT: Szabo, Veronique  
 TITLE OF INVENTION: Procedures and Materials for Confering  
 TITLE OF INVENTION: Disease Resistance in Plants  
 NUMBER OF SEQUENCES: 27  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/587,680A  
 FILING DATE: 17-JAN-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/373,375  
 FILING DATE: 17-JAN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/475,891  
 FILING DATE: 07-JUN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/004,645  
 FILING DATE: 29-SEP-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/567,375  
 FILING DATE: 04-DEC-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-058940US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3921 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: join(1..2676, 3520..3918)  
 OTHER INFORMATION: /product="A-21"  
 US-08-587-680A-3

	Query Match	1.18;	Score 46.2;	DB 4;	Length 3921;
	Best Local Similarity	49.48;	Pred No. 0.041;		
	Matches 120;	Conservative	0;	Mismatches 123;	Indels 0;
				Gaps	0;
QY	2114	acaaacaaatccaaagatttcagacagaagaagcactggaattacccttttglytaattgaatc	2173		
Db	1054	AAATTCAAAACATTGAACCTTGGAGAAATAATACCTGGGGGGAGTCTTCCATATTCGTT	1113		
QY	2174	tttttaacaagttctattttctctacagggagcttttcagtttaacacataactgycgcg	2233		
Db	1114	TCCATCTTTCACCTTCGCTTATCTTCTTTCCTGGACCTTGATTAAGATACACGAGAC	1173		
QY	2234	atttcctgtaattatttggaaatctcgacaaactcagtgagtttggacttctaactaaacgc	2293		
Db	1174	ATTCCGAGAGATATTGGCAATCTTATTGGCTTACCAACATCTCTATCTCTGCAACAAAT	1233		
QY	2294	ttctcgcgtcctattcccgaaatcattgggaaagcttcaaaagctgagatttctgtgta	2353		

Db	1234	TTCAGAGGTC	CTTC	CCATCATC	GTG	GGAGGCTT	AAAACTT	AGGCA	TTCT	ACTGCC	1233
Qy	2354	tac	2356								
Db	1294	TAC	1296								

```

14 RESULT
15 US-08-475-891A-3
16 Sequence 3, Application US/08475891A
17 Patent No. 5859339
18 GENERAL INFORMATION:
19 APPLICANT: Ronald, Pamela C.
20 APPLICANT: Wang, Guo-Liang
21 APPLICANT: Song, Wen-Yuang
22 TITLE OF INVENTION: Procedures and Materials for Confering
23 TITLE OF INVENTION: Disease Resistance in Plants
24 NUMBER OF SEQUENCES: 15
25 CORRESPONDENCE ADDRESS:
26 ADDRESSEE: Townsend and Townsend and Crew LLP
27 STREET: Two Embarcadero Center, Eighth Floor
28 CITY: San Francisco
29 STATE: California
30 COUNTRY: USA
31 ZIP: 94111-3834
32 COMPUTER READABLE FORM:
33 MEDIUM TYPE: floppy disk
34 COMPUTER: IBM PC compatible
35 OPERATING SYSTEM: PC-DOS/MS-DOS
36 SOFTWARE: PatentIn Release #1.0, Version #1.30
37 CURRENT APPLICATION DATA:
38 APPLICATION NUMBER: US/08/475,891A
39 FILING DATE: 06-JUN-1995
40 CLASSIFICATION: 800
41 PRIOR APPLICATION DATA:
42 PRIOR APPLICATION NUMBER: US 08/373,375
43 FILING DATE: 17-JAN-1995
44 ATTORNEY/AGENT INFORMATION:
45 NAME: Bastian, Kevin L.
46 REGISTRATION NUMBER: 34,774
47 REFERENCE/DOCKET NUMBER: 02370-058910US
48 TELECOMMUNICATION INFORMATION:
49 TELEPHONE: (415) 576-0200
50 TELEFAX: (415) 576-0300
51 INFORMATION FOR SEQ ID NO: 3:
52 SEQUENCE CHARACTERISTICS:
53 LENGTH: 5992 base pairs
54 TYPE: nucleic acid
55 STRANDEDNESS: single
56 TOPOLOGY: linear
57 MOLECULE TYPE: DNA (genomic)
58 FEATURE:
59 NAME/KEY: CDS
60 LOCATION: join(512..3149, 3993..4393)
61 OTHER INFORMATION: /product= "RRK-B"
62 OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
63 OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
64 sativa)"
65 US-08-475-891A-3

```

	Query Match	Similarity	1.1%	Score	46.2	DB	3	Length	5992
	Best Local	Similarity	49.4%	Pred.	No.	0.051			
	Matches	120	Conservative	0	Mismatches	123	Indels	0	Gaps
QY	2114	acaaacaaatccagatttcagagagagacacgagatcccttttgattgaatc	2173						
DB	1565	AAATTTCAAACATTGCACTTGGAGAAATTAACCTGGGGGGATTCTTCTTAATTCGTTT	1624						
QY	2174	ttttacaagttcttattttcttcacagggacgtttacagtaacacataacgtgcccg	2233						
DB	1625	TCCAACTCTTTCACCTTCGGTTAGTTTCTTCGACATTGAATTAAATAGATCAGAGAAC	1684						



QY 2234 attcctagtaacttgcgaacttgacaaacttagtgagtttggtacttacttaaacagc 2293  
|||||  
Db 1665 ATTCCGAAGGATATTGGCAATCTTATTGGCTTACAAACATCTATCTGCAACACACAT 1744  
QY 2294 ttctccggtctatttcgcgaactcattggaaagcttcaaaagctgagattctgtagta 2353  
|||||  
Db 1745 TTCAGAGGGTCTCTTCATCATCGTTGGCGACGCTTAAAACTTAGCATCTACTCGCC 1804  
QY 2354 tac 2356  
|||  
Db 1805 TAC 1807

RESULT 15  
US-08-475-891A-1  
; Sequence 1, Application US/08475891A  
; Patent No. 5859339  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; TITLE OF INVENTION: Procedures and Materials for Conferring  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,891A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 02370-058910US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6256 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1648..4383, 5178..5513)  
; OTHER INFORMATION: /product= "RRK-F"  
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease  
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza  
; OTHER INFORMATION: sativa)"  
US-08-475-891A-1

Query Match 1.1%; Score 46.2; DB 3; Length 6256;  
Best Local Similarity 49.4%; Pred. No. 0.052;  
Matches 120; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 2114 acaaacacatccagattgacagagaagcactgagttacttctgtaatgaatc 2173

Db 2722 AATTACAAACATGTGACTTGGGAGAAAATAACCTGGGGGAGTCTTCCATATCGTT 2781  
|||||  
QY 2174 ttcttaacaagtttcttatttcttaacaggagagcttcaagtaacaataactgagccg 2233  
|  
Db 2782 TCCAACTTTCACCTTCGCTGATTGTTCTTGCACTTGATTTGAATTAAGATCCAGGAGC 2841  
QY 2234 attcctagtaacttgcgaacttgacaaacttagtgagtttggaatttacttaaacagc 2293  
|||||  
Db 2842 ATTCCAAAGGATATTGGCAATCTTATTGGCTTACAAACATCTATCTCTGCAACACACAT 2901  
QY 2294 ttctccggtctatttcgcgaactcattggaaagcttcaaaagctgagattctgtagta 2353  
|||||  
Db 2902 TTCAGAGGGTCACTTCATCATCGTTGGCGACGCTTAGAAAATTAGCATTCAGTCGCC 2961  
QY 2354 tac 2356  
|||  
Db 2962 TAC 2964

Search completed: June 23, 2000, 22:44:24  
Job time: 40210 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:03 ; Search time 6198.48 Seconds  
(without alignments)  
2668.590 Million cell updates/sec

Title: US-09-180-798-20  
Perfect score: 4081  
Sequence: 1 tctagaaccttgcattcat.....tacggaatcatgcttctaga 4081

Scoring table: IDENTITY\_NTC  
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: em\_est10:\*  
11: em\_est11:\*  
12: em\_est12:\*  
13: em\_est13:\*  
14: em\_est14:\*  
15: em\_est15:\*  
16: em\_est16:\*  
17: em\_est17:\*  
18: em\_est18:\*  
19: em\_est19:\*  
20: gb\_est1:\*  
21: gb\_est2:\*  
22: gb\_est3:\*  
23: gb\_est4:\*  
24: gb\_est5:\*  
25: gb\_est6:\*  
26: gb\_est7:\*  
27: gb\_est8:\*  
28: gb\_est9:\*  
29: gb\_est10:\*  
30: gb\_est11:\*  
31: gb\_est12:\*  
32: gb\_est13:\*  
33: gb\_est14:\*  
34: gb\_est15:\*  
35: gb\_est16:\*  
36: gb\_est17:\*  
37: gb\_est18:\*  
38: gb\_est19:\*  
39: gb\_est20:\*  
40: gb\_est21:\*  
41: gb\_est22:\*  
42: gb\_est23:\*  
43: gb\_est24:\*  
44: gb\_est25:\*

45: gb\_est26:\*  
46: gb\_est27:\*  
47: gb\_est28:\*  
48: gb\_est29:\*  
49: gb\_est30:\*  
50: gb\_est31:\*  
51: gb\_est32:\*  
52: gb\_est33:\*  
53: gb\_est34:\*  
54: gb\_est35:\*  
55: gb\_est36:\*  
56: gb\_est37:\*  
57: gb\_est38:\*  
58: gb\_est39:\*  
59: gb\_est40:\*  
60: gb\_est41:\*  
61: gb\_est42:\*  
62: gb\_est43:\*  
63: gb\_est44:\*  
64: gb\_est45:\*  
65: gb\_est46:\*  
66: gb\_est47:\*  
67: gb\_est48:\*  
68: gb\_est49:\*  
69: gb\_est50:\*  
70: gb\_est51:\*  
71: gb\_est52:\*  
72: gb\_est53:\*  
73: gb\_est54:\*  
74: gb\_est55:\*  
75: gb\_est56:\*  
76: gb\_est57:\*  
77: gb\_est58:\*  
78: gb\_est59:\*  
79: gb\_est60:\*  
80: gb\_est61:\*  
81: gb\_est62:\*  
82: gb\_est63:\*  
83: gb\_est64:\*  
84: gb\_est65:\*  
85: gb\_est66:\*  
86: gb\_est67:\*  
87: gb\_est68:\*  
88: gb\_est69:\*  
89: gb\_est70:\*  
90: gb\_est71:\*  
91: gb\_est72:\*  
92: gb\_est73:\*  
93: gb\_est74:\*  
94: gb\_est75:\*  
95: gb\_est76:\*  
96: gb\_est77:\*  
97: gb\_est78:\*  
98: gb\_est79:\*  
99: gb\_est80:\*  
100: gb\_est81:\*  
101: gb\_est82:\*  
102: gb\_est83:\*  
103: gb\_est84:\*  
104: gb\_est85:\*  
105: gb\_est86:\*  
106: gb\_est87:\*  
107: gb\_est88:\*  
108: gb\_est89:\*  
109: gb\_est90:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES



```

Arabidopsis thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenomc.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 76
High quality sequence stop: 704.

FEATURES
      source
          1..1202
            /organism="Arabidopsis thaliana"
            /strain="Columbia"
            /db_xref="taxon:3702"
            /clone="T2E10"
            /clone_11b="TAMU"
            /sex="hermaphrodite"
            note="Vector: BelosACII, Site_1: HindIII, Site_2:
            HindIII; Produced by Rod Wing"

BASE COUNT      322 a      275 c      238 g      354 t      13 others

ORIGIN

Query Match      5.7%; Score 233.8; DB 84; Length 1202;
Best Local Similarity 66.6%; Pred. No. 4,4e-39;
Matches 428; Conservative 0; Mismatches 146; Indels 69; Gaps

QY 3002 aattgattccaggttgatggtatgtaactagcagcaatagctgtgtgagttgctgcaggt 3061
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 AATTAATTTAAAGTGGATAGTCTCTACTGAGCAGCATGCGGAGAGAGTGTCTGTGT 187

QY 3062 gctgtgtgctcttcgtcgtccctcgcgaatagcccttgcgtgtgtgcgacgaagaacca 3121
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 GCTGCTTACTATTCTGCTCCCTCCTTACGCTTTTGCTTGCTGGTGGCTAAAAAAACCT 247

QY 3122 ctgagatcttcgcatgtgcctcgtgagtttatctatctgcattagttctctcttag 3181
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 CAGAAATTCCTCTTATGATGTCTCTGGTAGTCACCT-----GAGTCTGCATATTC 296

QY 3182 ccagccaatttgcgttcgcgaagaagatctggaacacatgtaatgaaatcaatacataa 3241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 CAAGCTTGTGTTCAATTTCAGAAATGGAAT-----A 326

QY 3242 gtcaattgttttaagattcaaacctctttgaagttaaatctcgattgcaaatctctag 3301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

RESULT 2
LOCUS   AQ969360      714 bp    DNA
DEFINITION LERM537F LERG Arabidopsis thaliana genomic clone LERM53, genomic survey sequence.
ACCESSION AQ969360
VERSION  AQ969360.1  GI:6797061
KEYWORDS GSS.
SOURCE   thale cress.
ORGANISM Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
           eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
           Arabidopsi.
REFERENCE 1 (bases 1 to 714)
           Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T.,
           Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.
           Genomic survey sequencing of Landsberg erecta ecotype of
           Arabidopsis thaliana and identification of sequence-based
           polymorphisms
JOURNAL   Unpublished (2000)
COMMENT   On Dec 15, 1999 this sequence version replaced gi:4215892.
           Contact: Xiaoying Lin
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: at@tigr.org
           For additional information, see http://www.tigr.org/tdb/at/at.html
           Seq primer: TF
           Class: shotgun.
FEATURES
   source          Location/Qualifiers
                   1..714
                   /organism="Arabidopsis thaliana"
                   /strain="Landsberg erecta"
                   /db_xref="taxon:3702"
                   /clone_1ib="LERM53"
                   /note="Organ: leaf; Vector: pUC19UK; Total genomic DNA was
                   sheared to 0.4-0.7 Kbp before ligation."
BASE COUNT  205 a      177 c      112 g      220 t
ORIGIN
Query Match      5.2%; Score 213; DB 92; Length 714;
Best Local Similarity 60.9%; Pred. No. 1e-34;
Matches 389; Conservative 0; Mismatches 210; Indels 40; Gaps 1;
QY 3443 cagacggaactctgtgtcgtcgaagactgaagagagcgaactcgaagtggagagc 3502
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 702 CGATGGCACCACCAAAATCGCTGTAAGAAAGATTGACTGATTGAACTCCAGAGGAGANG 643
QY 3503 tccagttcaaacgaagatgagatgataatgtagcagttcgcgaactctgtaagat 3562
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 642 AAGCTTCCAGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 583
QY 3563 taagaagttctgtatgacacgacgagagattgtgtgtatccttaccatgagccaag 3622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 582 TTATCGGCTTTGTGACACACAACTGAAGACTTTGTGATCTTTCATTCACAGAAATC 523
QY 3623 gaagtgttgcgtcgtcctcagaggtaaaactaaacaattaaacactctgtgctctc 3682
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 522 TAACTGTTGCATATTTGCTTAAAGAGTAAAGAAAGAACATNA----- 480
QY 3683 tcaattacttgagtgagtgatgttttctgtttcctttagtggttcaattgttg 3742
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 480 -----GCTATTAACCTTTTGTATTTCAAGAGATTTAAGATTTTGTG 443
QY 3743 ttacactaatgacacagagagggccacccgtcacaacctcgcgttgattggccaagcgagaa 3802
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 442 ATCTATATCTATGTCAGAGATTAAACCCGGGATCCAGTTCTGATTTGTTCAAGAGAA 383

```

```

QY 3803 gagaatcgcgtagtcgctacgctcgcgcgcgtgttgccttacctacatgctacgtccgaa 3862
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 382 ACAGATTGCTGTAGTGACAGACGAGACTCGAATATCTTCATGAACTATGCAACCCGAA 323
QY 3863 gatcatcaccgtgagtaaaacagacaacacatccctttagacgagaattcgaagcgt 3922
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 322 GATCATACACAGAGATGTGAAGCTCAAAATGTTTACTAGATGAAGACTTTGAAGCAGT 263
QY 3923 tgttgagatctcgggttggaagactaaatgaactataaagaactcagtgacacagc 3982
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 262 GGTGGGATTTGTTGTTAGCCAGTTGGTAGATGTTTGAAGACTAATGTGAACCACTCA 203
QY 3983 agtccgtgaccatcgatgcacatgcctcagaatactcctcaacgggaaacttcaga 4042
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 202 GGTCCAGAGAACATGGGTATTTTCACACAGATGTATATTCACAGAGGAATTCGTGGA 143
QY 4043 gaaacgcagcttttcgagatcgaatcattcgtctcaga 4081
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 142 GAAACCGATGTTTTCGGGTGCGGATGATGCTTGTGGA 104
RESULT 3
LOCUS   AQ969361      621 bp    DNA
DEFINITION LERM537R LERG Arabidopsis thaliana genomic clone LERM53, genomic survey sequence.
ACCESSION AQ969361
VERSION  AQ969361.1  GI:6797062
KEYWORDS GSS.
SOURCE   thale cress.
ORGANISM Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
           eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
           Arabidopsi.
REFERENCE 1 (bases 1 to 621)
           Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T.,
           Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.
           Genomic survey sequencing of Landsberg erecta ecotype of
           Arabidopsis thaliana and identification of sequence-based
           polymorphisms
JOURNAL   Unpublished (2000)
COMMENT   On Dec 15, 1999 this sequence version replaced gi:4215893.
           Contact: Xiaoying Lin
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: at@tigr.org
           For additional information, see http://www.tigr.org/tdb/at/at.html
           Seq primer: TR
           Class: shotgun.
FEATURES
   source          Location/Qualifiers
                   1..621
                   /organism="Arabidopsis thaliana"
                   /strain="Landsberg erecta"
                   /db_xref="taxon:3702"
                   /clone_1ib="LERM53"
                   /note="Organ: leaf; Vector: pUC19UK; Total genomic DNA was
                   sheared to 0.4-0.7 Kbp before ligation."
BASE COUNT  191 a      102 c      154 g      173 t
ORIGIN
Query Match      5.2%; Score 210.4; DB 92; Length 621;
Best Local Similarity 60.6%; Pred. No. 3.5e-34;
Matches 387; Conservative 0; Mismatches 212; Indels 40; Gaps 1;
QY 3443 cagacggaactctgtgtcgtcgaagactgaagagagcgaactcgaagtggagagc 3502
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 21 CGATGGCACCACCAAAATCGCTGTAAGAAAGATTGACTGATTGAACTCCAGAGGAGANG 80

```



[illegible]

RESULT	4
LOCUS	AM233982
DEFINITION	AM233982 532 bp mRNA EST 13-DEC-1999
ACCESSION	G532905.y1
VERSION	gm-cl028
KEYWORDS	glycine max cDNA clone, GENOME SYSTEMS CLONE ID
SOURCE	gm-cl028-1/13.5, similar to tr:023921 023921 SOMATIC EMBRYOGENESIS
ORGANISM	RECEPTOR-LIKE KINASE. ; mRNA sequence.
	AM233982
	AM233982.1 GI:65636309
	EST.
	soybean.
	glycine max
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	euphyllophytes; Spermatophytes; Magnoliophyta; eudicotyledons; core
	eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
	Glycine.
REFERENCE	1 (bases 1 to 532)
AUTHORS	Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Coryell, V., Khanam, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wille, T., Underwood, R., Steptoe, M., Theising, B., Allen, M., Bowers, J., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	Public Soybean EST Project
JOURNAL	Public Soybean EST Project
COMMENT	Unpublished (1999) On Feb 18, 1999 this sequence version replaced g1:4297707. Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu

FEATURES	source
1.. 532	location/Qualifiers
..organism="Glycine max"	
/db_xref="taxon:3847"	
/clone="GENOME SYSTEMS CLONE ID: Gm-cl028-1713"	
/clone_lib="Gm-cl028"	
/tissue_type="roots of 'Superpod' plants"	
/lab_host="DH10B"	
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of Glycine max 'Superpod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicum, strain USDA110 prior to harvest. Stratagene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5'-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGACTGCTGCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Gencobri Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryelli."	
BASE COUNT	145 a 92 c 129 g 165 t 1 others
ORIGIN	
Query Match	4.7%; Score 192.8; DB 79; Length 532;
Best Local Similarity	74.7%; Pred. No. 1.7e-30;
Matches 242; Conservative 0; Mismatches 82; Indels 0; Gaps 0;	
Oy 3758 agaaagcgacacgtcacacacgtcgcttgatggccaaagcgaaagaaatcgctagg 3817	
Db 103 AGAAGCGTCAAGAAATCCCAACCGCCACTTGCGTGGCCAGAACGACGATTTGCATTGGG 162	
Oy 3818 ctcaagctcgaaggttcttaccatcatgatcactgcgaticgaticcgaaagatcatcaocgtga 3877	
Db 163 ATCTGCAAGGGGCTTGCTTATTTCAGATGATCATTTGACCTCAAGATATTCACCGCTGA 222	
Oy 3878 cgtaaagcgacgaacatcctcttgaagaaattcgaagcggttttttgagatttgg 3937	
Db 223 TGTGAAAGCGACGTAAATATTTTGTGGAGAGAGAAATTTGAAGCACTGTTTGGAAATTTTGG 282	
Oy 3938 gttgcgaagctcatcgtacataaagacactcgcgtgacaaacgacgttcgtgcacat 3997	
Db 283 TTTAGCTTAACTTATGATTAATAAAGATACATCACTACTACGCTGCTATGTAACAAT 342	
Oy 3998 cagtgacatcgtcgcgaatcatctcgaacggaatcttcgagaaacggaagctttt 4057	
Db 343 TGGACATATATGCAACGAAATACCTCTTCACACTGGAAAGTCTTCAAGAAAGACTGATGTTTT 402	
Oy 4058 cgatatacgaaatcatgcttctaga 4081	



Db	403	TGGAATATGCTGTGATGCTTTCTTGA	426
RESULT	5		
LOCUS	A1486453		
DEFINITION	A1486453 471 bp mRNA EST 29-JUN-1999		
ACCESSION	EST244774	tomato ovary, TAMU Lycopersicon esculentum cDNA clone	
VERSION	CL24882	mRNA sequence.	
KEYWORDS	A1486453.1	GI:4381824	
SOURCE	EST.		
ORGANISM	tomato.		
REFERENCE	Lycopersicon esculentum		
AUTHORS	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.		
TITLE	1 (bases 1 to 471)		
JOURNAL	Alcala, J., Vrebalov, J., White, R., Matern, A.L., Viston, T., Holt, L.E., Liang, F., Opton, J., Konning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Niernan, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.		
COMMENT	Generation of ESTs from tomato carpel tissue Unpublished (1999)		
	On May 7, 1998 this sequence version replaced gi:3121325.		
	Contact: David Frisch		
	Clemson University Genomics Institute		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Tel: 864 656 4366		
	Fax: 864 656 4283		
	Email: dfrisch@CLEMSON.EDU.		
FEATURES	Location/Qualifiers		
SOURCE	1..471		
	/organism="Lycopersicon esculentum"		
	/cultivar="TA96"		
	/db_xref="taxon:4081"		
	/clone="CL24882"		
	/clone_id="tomato ovary, TAMU"		
	/tissue_type="carpel"		
	/dev_stage="5 days pre-anthesis to 5 days post-anthesis"		
	/lab_host="X11-Blue MR"		
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and directionally cloned cDNA in vector lambda ZAP II with 5' and 3' ends located at the EcoRI and XhoI sites, respectively"		
BASE COUNT	136 a 88 c 112 g 135 t		
ORIGIN			
Query Match	4.2%; Score 169.4; DB 47; Length 471;		
Best Local Similarity	72.0%; Pred. NO. 1.4e-25;		
Matches 221; Conservative	0; Mismatches 86; Indels 0; Gaps 0;		
QY	3775 aactcgcgttgatgagccaacgcygaagagatcgcgctagagctcgaagtttgt	3834	
Db	70 AGCGTGTGTAATGGCCAACTAGAGACGCGTGCGGTGGTACTGACGTGACTTG	129	
QY	3835 cttaactacatgatactacgtgcgataccgaagatcattcaaccgtgacgttaaaacagcaaca	3894	
Db	130 AATACCTCATGTAACACGTAAATCCAAAGATTAATTCACCGGTATGTAAAGCAGCTAATG	189	
QY	3895 tccctctgagcaagaattcgagacgagtggtttgagagattcggttggcaagcctaagtg	3954	
Db	190 TGTACTAGATCAAAATTTTGAAGCTGTAGTTGGTGATTTTGGCCTGGCAAGCTAGATTG	249	
QY	3955 actataaagacactcagctgacaaacagcagtcgctggcaccatcggctacatcgcgtcag	4014	
Db	250 ACGTTAAGAAACCAATGAGCTACTCAAGTTCGTGGTCAATGGGCCATATAGTCTCG	309	
QY	4015 aatactctcaacgcggaatacttcaagagaaacaggaagtttcggatacggatcatcag	4074	

Db	310	AAATCTATTCACATGGCAAAATCATCAGAAAAAATACTGATGTTTGGCTATGAAATCATGC	369
QY	4075	ttctaga 4081	
Db	370	TTTTGGA 376	
RESULT	6		
LOCUS	AA738544	991 bp mRNA	EST 17-AUG-1998
DEFINITION	SbRLK2 Sorghum bicolor cv. TX430 leaf Sorghum bicolor cDNA clone		
ACCESSION	AA738544		
VERSION	AA738544.1	GI:3421469	
KEYWORDS	EST.		
SOURCE	Sorghum.		
ORGANISM	Sorghum bicolor		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.		
TITLE	1 (bases 1 to 991)		
AUTHORS	Annen, F., Chang, J.-I., Paterson, A.H. and Stockhaus, J.		
JOURNAL	Characterization of 14 different putative protein kinase cDNA clones of the C4 plant Sorghum bicolor		
MEDLINE	Mol. Gen. Genet. 259 (1), 115-122 (1998)		
COMMENT	98409267 On Nov 29, 1993 this sequence version replaced gi:636080. Contact: Stockhaus J Institut fuer Entwicklungsbiologie und Molekularbiologie der Pflanzen Heinrich-Heine-Universitaet Universitaetsstrasse 1, 40225 Duesseldorf, Germany Tel: (49)-211-81-14953 Fax: (49)-211-81-14871 Email: stockhaeuni-duesseldorf.de		
FEATURES	Insert length: 1500 Std Error: 0.00.		
source	Location/Qualifiers 1..991 /organism="Sorghum bicolor" /strain="cv. TX430" /db_xref="taxon:4558" /clone="SbRLK2" /clone_id="Sorghum bicolor cv. TX430 leaf" /dev_stage="green" /note="Organ: leaf; Vector: Lambda ZAP II; Site.1: Eco RI; Site.2: Xho I; leaf cDNA library from green leaves, unidirectionally cloned"		
BASE COUNT	253 a 194 c 293 g 251 t		
ORIGIN			
Query Match	4.1%; Score 167; DB 38; Length 991;		
Best Local Similarity	55.7%; Pred. No. 4.0e-25;		
Matches 428; Conservative 0; Mismatches 225; Indels 116; Gaps 1;			
Db	3313	gattccagaagttcatctcttggagacgtctcaagggtttcttcttggggagagctcaagtgagc 3372	
QY	1	GATCGTAGAATTGCAATTTGGCCAGTTGAAAAGATTTCATGAGAAAGATTGCCAATCTTGA 60	
Db	3373	agtagaggttagtatacaagaacatttggcgagagtgaggttgggaaagctacaag 3432	
QY	61	ACGATGATTAATTTCAGTGTAGAAAAAATGTTCTTGAGACAAGGGGCTTTGGGAAAGTATTTAA 120	
Db	3433	ggaagccttggcagaaggaactcttcttgcgtctcaagagactgaaggaagagcgaaactca 3492	
Db	121	GGAGCACTTCACAGATGGCACTAGATTCTCTGTAAAACGGTTAACTGATTACGAAGTCTT 180	
QY	3493	gagtgagagagctccagtttcaaacagaatagaagatgataagtatgagcagttcctgaac 3552	
Db	181	GGTGGAGAGGCTGCTTTCTTGGTGAAGGTTAGAGTGATATGATGTGGAGTTACACCGGAAT 240	



QY	3553	ctgttgagatgaagaggtttctctatgaacacgagacgaagatggtctgtatccctac	3612
Dd	241	CTTTAAAGATTGATGGTGTTCCTGTACACACCAACAGAGCGCTGTTTATATCTTTC	300
QY	3613	atggccaaatggaagtgctgtctcgtgclctcagaagtataaactaaacaatlaaacactt	3672
Dd	301	ATGCACATCTTATGTTT-----	319
QY	3673	gtgtctctctccaaacttaacttttgatcgaagtggtttttcatgttttcccttatggttca	3732
Dd	319	-----	319
QY	3733	taatttgtttacactatgacacagagagcgccacgcgtcaaacctccgtcttgatgc	3792
Dd	319	-----TCTACCGTCTACAGAGATTTAACCTGGGGACCAATATTAGATTGGT	364
QY	3793	caacgggaagagatctgcgtctgctcagcttgaaagttgtcttaaccactatgatact	3852
Dd	365	CTGCAAGGAAGCGATGGCTATTAGGCAACACTGTGGACTGGAGATTTTGGACGAGCACT	424
QY	3853	gcgactccgaagatcatcaccgcgtgacgttaaaagcaacaacatccctcttagacgaagat	3912
Dd	425	GCAATCCTAAGATTATACATCGTAGAGCTCAAGGCTGCCAATGTCTGTATGAAGGTT	484
QY	3913	tcgaagcgggtgtgttggagatcttcggggttggcaagcttaatgtaacttaagaacactaacg	3972
Dd	485	TTTGAAACGGTTGGTGGATTGGCTTGCGCTGGCCAAAGCTGGTGATGACGAAGAAACATCTNG	544
QY	3973	tgaaacacagcagtcctgcgtgcacatcagttacactcgtccagataatctctcaacccgaa	4032
Dd	545	TGACTACTCAGAGTCCGATGGANCAATAGGATCAATTCGCCCTGAAGATTGTCCACATGGGA	604
QY	4033	aactctcagagaacacgcagcgttttcggatacgaagatcatgctctctaga	4081
Dd	605	AGTATCCGAGAGAACCGATGTTTGGTTACGCGATTAATGCTTTCGA	653

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 475)	Harrison,M.J., Liu,J., Gonzales,M.B. and Ellis,L.	ESTs from phosphate starved roots	Unpublished (1999)	
		On May 1, 1997 this sequence version replaced gi:2059389.		
	Contact: Maria J. Harrison			
	Plant Biology Division			
	The Samuel Roberts Noble Foundation			
	2510 Sam Noble Parkway, Ardmore, OK 73401, USA			
	Tel: 580-223-5610			
	Fax: 580-221-7380			
	Email: mjharrison@noble.org			
	Developmental stages: phosphate starved; Date: 01/23/00; Submitted			
	to the database of Expressed Sequence Tags (dbEST) on 01/26/00;			
	More information is available at			
	'http://chryslc.tamu.edu/medicago'.			
	Seq primer: T3.			

```

FEATURES
source      location/Qualifiers
1. .475
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"

```

```

/clone="MHRP-24H06"
/clone.lib="rootphos(-)"
/tissue_type="roots"
/note="Vector: pBluescript SK -; Site.1: EcoRI, Site.2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The
cDNA was directionally ligated into the unzip XR vector
from Stratagene and packaged using Gigapack III Gold
packaging extracts. Plasmids containing cDNA inserts
were excised from the recombinant lambda-zap phase using
Ex-assist helper phage and propagated in SOLR cells."

```

Query Match	4.1%	Score 166.2	DB 80	Length 475
Best Local Similarity	71.3%	Pred. No. 6.7e-25		
Matches 219	Conservative	0	Mismatches 86	Indels 0
				Gaps 0
Qy 3775	aactccgctgtgattgycacagcggaaagaatcgcgtlaagcccaagctcgattgt	3834		
Db 68	AGCCAGCTTTAGACTGGGCACACAGGAGAGAGATAGCATTAGAGCAGGAGAGGGTTAC	127		
Qy 3835	cttaaccatgatcactcgcgatccgaagatcatcaccgtgaaglaaaagcagaaca	3894		
Db 128	TATATTTACATGACACACTGTGATCCAAAGATATATTCACAGAGATGTTAAAGCAGCAATA	187		
Qy 3895	tcacctgaagcagaatttcgaagcgtgtgttgagattcggggttgcaagcctaaag	3954		
Db 188	TATGTCTAGATGATTATGTGAGGCTGTGTGTGGAGACTTGTGGATTGCGAAAATCTATTGG	247		
Qy 3955	actataagaacactcactgtacacacagcagctccgtgtgcacactcgttcacatcgtccag	4014		
Db 248	ACCAACAGAGACTCACAATGTACAAACAGCAGTGAGAGACTTAGAGACACATAGCCCTG	307		
Qy 4015	aatatctctcaacccggaaaatcttcagagaaaaacggagcttccgatacggaaatcagc	4074		
Db 308	AATACCTATCAACAGGACCAATCCTCAGAGAAAACAGATATTGTTGGAATTTCTC	367		
Qy 4075	ttctaga 4081			
Db 368	TGCTTGA 374			

RESULT	8
LOCUS	AA738547
DEFINITION	AA738547 656 bp mRNA EST 17-AUG-1998
ACCESSION	SBELK5 Sorghum bicolor cv. TX30 leaf Sorghum bicolor cDNA clone
VERSION	SBELK5 similar to protein kinase, highest similarity to
KEYWORDS	receptor-like protein kinases, mRNA sequence.
SOURCE	AA738547
ORGANISM	AA738547.1 GI:3421472
	EST.
	sorghum.
	Sorghum bicolor
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	euphylliphytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
	Poaceae; Sorghum.
REFERENCE	1 (bases 1 to 656)
AUTHORS	Amenu, F., Chang, J., -L., Paterson, A.H. and Stockhaus, J.
TITLE	Characterization of 14 different putative protein kinase cDNA
JOURNAL	Clones of the C4 plant Sorghum bicolor
COMMENT	Mol. Gen. Genet. 259 (1), 115-122 (1998)
	98409267
	On Nov 23, 1993 this sequence version replaced gi:636097.

Institut fuer Entwicklungsbiologie und Molekularbiologie der  
Pflanzen  
Heinrich-Heine-Universitaet  
Universitaetsstrasse 1, 40225 Duesseldorf, Germany  
Tel: (49)-211-81-14953  
Fax: (49)-211-81-14871  
Email: stoeckhaun-duesseldorf.de







LOCUS	25047	299 bp	mRNA	EST	08-JUL-1999
DEFINITION	R1C92976	Rice root	Oryza sativa	cdna	clone R2976_1A, mRNA
ACCESSION	D25047				
VERSION	D25047.1	GI:428891			
KEYWORDS	EST.				
SOURCE	Oryza sativa.				
ORGANISM	Oryza sativa				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.				
AUTHORS	1 (bases 1 to 299)				
TITLE	Minobe, Y. and Sasaki, T.				
JOURNAL	Rice cdna from root				
COMMENT	Unpublished (1995)				
	Contact: Yuzo Minobe				
	National Institute of Agrobiological Resources				
	Rice Genome Research Program				
	2-1-2 Kannondai, Tsukuba				
	Ibaraki,				
	Japan 305				
	Tel: 0298-38-7441				
	Fax: 0298-38-7468				
	Email: minobe@rcs0.riken.go.jp				
FEATURES	PROJECT "Rep".				
source	Location/Qualifiers				
	1..299				
	/organism="Oryza sativa"				
	/strain="Nipponbare, sub-species japonica"				
	/db_xref="taxon:4530"				
	/clone="R2976_1A"				
	/clone_1lb="Rice root"				
	/note="Prepared from seedling root."				
BASE COUNT	93 a	54 c	67 g	85 t	
ORIGIN					
Query Match	4.08;	Score 164.4;	DB 20;	Length 299;	
Best Local Similarity	74.58;	Pred. No. 1.6e-24;			
Matches 207;	Conservative	0;	Mismatches 71;	Indels	0; Gaps
3804	agaatcgcgtaagcgtacgtcaggttgccttacctacatgatcgtcgtacgcgaag	3863			
Db	3 AGGATCCGACAGAGTTTGCCACGAGGGGCTCTCTATTACATGATCATTCGCAACCAAG	62			
Qy	3864 atcattacgctgacgtaaaagacgaacacatcctctttagacgaagaattcgaa	3923			
Db	63 ATCATCATGCTGATGTCNAAGCGCAATATTTATTGATGAGAGACTTTGAACTGTA	122			
Qy	3924 gttagcatttcggtgttggaagaacgaatgacataaagaacatcgtgacaaacga	3983			
Db	123 TTTCGAGTTTGTGTTTGCCCAACCAATGATTCAGAGATTCACCATTTAACTGCA	182			
Qy	3984 gtccgtgacacatcgttcacatcgtccacgaatctctcaaccggaatcttcagag	4043			
Db	183 GTACATGGACAAATTGGACATATTGCACAGATATCTTTCAACAGAGATTCCTCTGAG	242			
Qy	4044 aaacgcagcttccgatacgaatcatgctctctga	4081			
Db	243 AAAACTGATGATTTGGCTATGAGATCATCTTTTGA	280			
RESULT	11				
LOCUS	B27171	464 bp	DNA	GSS	13-OCT-1997
DEFINITION	T2L10TF TAMD Arabidopsis thaliana genomic clone T2L10, genomic survey sequence.				
ACCESSION	B27171				
VERSION	B27171.1	GI:2513137			
KEYWORDS	GSS.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				

REFERENCE	TITLE	JOURNAL	COMMENT
1 (bases 1 to 464)	Rounsley, S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and Venter, J.C.	Use of a BAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing	Unpublished (1997)
Other_GSSs: T2E10TR	Contact: Steve Rounsley	Department of Eukaryotic Genomics	The Institute for Genomic Research
		9712 Medical Center Dr., Rockville, MD 20850, USA	Tel: 301 838 0200
			Fax: 301 838 0208
			Email: rounsley@tigr.org
			Seq primer: M13-21
			Class: BAC ends
			High quality sequence stop: 464.
FEATURES	source	location/Qualifiers	
	1..464	/organism="Arabidopsis thaliana"	
		/strain="Columbia"	
		/db_xref="taxon:3702"	
		/clone="T2E10"	
		/clone_11b="TAMU"	
		/sex="hermaphrodite"	
		/note="Vector: pBelbacII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"	
BASE COUNT	119 a	87 c	115 g 143 t
ORIGIN			
Query Match	3.9%;	Score 157.2;	DB 84; Length 464;
Best Local Similarity	64.5%;	Pred. No. 5.2e-23;	
Matches 320;	Conservative 0;	Mismatches 108;	Indels 68; Gaps 3;
QY 3002	aatttgatttcaggtggtatgtatgaacctggagcaatagctggtgagttgctgcaggt	3061	
DB 37	AATTAATTTAAAGTGGGTATAGTCTCTAGGACCAATGGGAGAGAGTTCCTGTGTG	96	
QY 3062	gctgccttgctcttcgctcctcctgcaatagaccttggctgtggtggcgaagaagaac	3121	
DB 97	GCTGCTTACTATTTCTGCCCCCTGCTTACGTTTCTGTTGCTGTGGCTAAGAAACCT	156	
QY 3122	ctagatatctcttcgcagatgctgcgtgagttattatttcgcattagtttcgttcttag	3181	
DB 157	CAAGAAATCTCTTTGATGTCTCTGTGAAGTACAT	205	
QY 3182	ccagcaatttggcttgcgagaaagatattggaacacactgtaatgaaatcacaatac	3241	
DB 206	CAACCTTGTGTTATTTCAATTCAGAAATGGATT	235	
QY 3242	gtcatgttttttaagtacaacaccttbtgataaatactcgattcaaaatctctatg	3301	
DB 236	GTAATTGACTCTTAATTTTTCAGTTCTGTG	268	
QY 3302	cagccgaagaagatccagaagtcatctcgtggagacagctcaagaaggtttcttcg	3361	
DB 269	CAGCCGAAGAGACCTGTAGGTTCACTTGGGGCACCTTAAACCGGTTCTCTACGGGAAC	338	
QY 3362	tacaaagtggagtgatggtttagttaacaagaacaatttgggcagaggtgggtttggga	3421	
DB 329	TTCAAGTAGAGACTGATAGCTTCAGCAACAGAAATTTTGGCCGAGGTGGGTTCGGA	388	
QY 3422	aagcttcaaaaggacgcttggcagacagcaactctgtgtcgttcaagaagactgaaggaag	3481	
DB 389	AAGCTCAACAATCCGCTGTGATGAGAACCTGTGTGCAGTCAACGCGCTTAAGGAAG	448	
QY 3482	agcgaactcaggtgg 3497		
DB 449	AACGAAACCCAAAGTGG 464		



RESULT	12
LOCUS	AA738545
DEFINITION	AA738545 1435 bp mRNA EST 17-AUG-1998 SRLK3 Scirpus bicolor cv. TX430 leaf Scirpus bicolor cDNA clone SRLK3 similar to protein kinase, highest similarity to receptor-like protein kinases, mRNA sequence.
ACCESSION	AA738545
VERSION	AA738545.1 GI:3421470
KEYWORDS	EST.
SOURCE	Scirpus. Scirpus bicolor
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Scirpum. 1 (bases 1 to 1435)
REFERENCE	Annen,F., Chang,J.-L., Paterson,A.H. and Stockhaus,J. Characterization of 14 different putative protein kinase cDNAs clones of the C4 plant Scirpus bicolor
AUTHORS	Mol. Gen. Genet. 259 (1), 115-122 (1998)
TITLE	
JOURNAL	
MEDLINE	98409267
COMMENT	On Nov 29, 1993 this sequence version replaced gi:636085.

FEATURES	SOURCE	Location/Qualifiers
		1. 1435
		/organism="Sorghum bicolor"
		/strain="cv. TX430"
		/db_xref="taxon:4558"
		/clone="SbRLK3"
		/clone_1lb="Sorghum bicolor cv. TX430 leaf"
		/dev_stage="green"
		/note="Organ: leaf; Vector: Lambda ZAP II; Site_1: Eco RI, Site_2: Xho I; leaf cDNA library from green leaves, unidirectionally cloned"
BASE COUNT		365 a 300 c 389 g 381 t
ORIGIN		

Query Match	3.9%	Score 157.2	DB 38	Length 1435
Best Local Similarity	70.5%	Pred. No. 5.3e-23		
Matches 210:	Conservative 0	Mismatches 88	Indels 0	Gaps 0
QY	3784	ttgatgtgccaa	cgcggaagaagatcgcgcgtagctcaagctcgaggttgccttaccac	3843
Db	800	TTGATTGGCTCTCA	GGAAGCGAGTGCGCTATAGGCGCACACTCGTGGACACGGATTATTTC	859
QY	3844	atgatcaatcgcgcgcgaagatcattaccgcgcgcgttaaacagagaacatccctttag	3903	
Db	860	ACGAGACCTGCACATCCTA	GATTATATCATCGTGACCTCAAGGCTGCCAATGCTTCCCTTG	919
QY	3904	acgaagaattcgaacgcggttcttggaagattcggggttgcgaagcctaattgactataaag	3963	
Db	920	ATGAAGGTTTTTAA	CCGGTGTGTGGATTGGGCTTGCCCAAGCTGGTGATGTACAA	979
QY	3964	aacatcaactgcgaacagacagtcgcgtgcgcacatcgtgtacatcgctccaagatartct	4023	
Db	980	AACAATCTGTGGCTGCTC	AGTCCCGTGGAACATATGGGTCAATATGCCCTCGAATATTATTTGT	1039
QY	4024	caacgcggaatatcttcagagaanaaccgcgcgttttcggatacggaaatcargctctaga	4081	
Db	1040	CCACGTGGGAAGTCATCC	GAGAGAACCGAATGTTTGGTTAACGCCAATAAAGCTTCTCGA	1097

RESULT	13
LOCUS	A1896277
DEFINITION	A1896277 555 bp mRNA EST 27-JUL-1999
ACCESSION	EST285720 tomato callus, TAMU Lycopersicon esculentum cDNA clone
VERSION	CLEC14J5, mRNA sequence.
KEYWORDS	A1896277 A1896277.1 GI:5602179
SOURCE	EST.
ORGANISM	tomato. Lycopersicon esculentum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyt euphyllaphytes; Spermatophyta; Magnoliophyta, eudicotyledons, co eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanu Lycopersicon. 1 (bases 1 to 555) Alcala, S., Vredalov, J., White, R., Matern, A.L., Vision, T., Holt, L.E., Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S., Ronnberg, C.M., Fraser, C.M., Martin, G.B., Tankeley, S.D. and Giovannoni, J. Generation of ESTs from tomato callus tissue Unpublished (1999) On May 18, 1998 this sequence version replaced gi:3137451.
TITLE	JOURNAL
COMMENT	

Email: [dfrisch@CLEMSON.EDU](mailto:dfrisch@CLEMSON.EDU)  
5 prime sequence.

```

FEATURES
source
    Location/Qualifiers
        1..585
        /organism="Lycopersicon esculentum"
        /cultivar="RA486"
        /db_xref="taxon:4081"
        /clone="cLEc14J5"
        /clone_lib="Tomato callus, TAMC"
        /tissue_type="callus"
        /dev_stage="25-40 days old"
        /lab_host="X11-Blue MRF"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; supplier: Giovannoni laboratory; cLEc - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST library"
BASE COUNT
154 a 90 c 148 g 163 t
ORIGIN

```

	Query Match	3.7%	Score 151;	DB 62;	Length 555;
	Best Local Similarity	71.3%	Pred. NO. 1e-21;		
	Matches 199;	Conservative	0;	Mismatches 80;	Indels 0;
				Gaps	0;
QY	3784	ttgatgtgcacacgcgggaagaatcgcgtacgtctcagctcgsagtttcttaacctac	3843		
DB	275	TTGGCTGTGTAAAGCGCGAAAGGAAATAGCGAGTGGGTACGACGAGGCGGCTGTATTATTGC	334		
QY	3844	atgacactcgcgtccgaaagatcattccacgcgagctaaagaacgaacatcctcttag	3903		
DB	335	ATGAAACAATGTGACCCCAAAATTATTCACCGTGAGTGTAAGCGACCAATTCGTGTGG	394		
QY	3904	acgaagaattcgaagcgggttgttggaattcgggttggcgaagctaaatgactataaag	3963		
DB	395	ATGAGGATTTGGAAGCAGTTGCGATTGTGGCTTTGCGCAAACTCTTGATACCGGG	454		
QY	3964	acaatcacgtgaaacaaacagcagtcctcgtggacacatcgtgacatcgcgcgaatatctc	4023		
DB	455	ATTCTCATGTAAAGACATCGCTGTGAGGGGCGACAGTTGGTCACATTTGCTCCAAAATATCTTT	514		
QY	4024	caaccggaaaaattcttcagagaaacgcgacgctttccgat	4062		
DB	515	CAACGGGTCAATCGTCTGAAAAGACGTGAGTGTGGT	553		



RESULT	14
LOCUS	A1727837
DEFINITION	BNLGH12209 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase [Daucus carota], mRNA sequence.
ACCESSION	A1727837
VERSION	A1727837.1 GI:5046689
KEYWORDS	EST
SOURCE	upland cotton.
ORGANISM	Gossypium hirsutum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; euclcotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
AUTHORS	Blawieit,M., Matz,E.C., Davy,D.F. and Burr,B.
TITLE	ESTs from developing cotton fiber
JOURNAL	Unpublished (1999)
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3187180. Contact: Ben Burr Biology Department Brookhaven National Laboratory Upton, NY 11973, USA Tel: 516-344-3396 Fax: 516-344-3407 Email: burrbnlx1.dn1.gov Seq primer: r3 primer.
FEATURES	Location/Qualifiers
source	1..597
	/organism="Gossypium hirsutum"
	/cultivar="Acala Maxxa"
	/db_xref="taxon:3635"
	/clone.lib="Six-day Cotton fiber"
	/tissue.type="Immature fiber"
	/dev.stage="Six days post anthesis"
	/lab.host="XLI-Blue"
	/note="Vector: pBluescript II KS+"
BASE COUNT	165 a 107 c 165 g 148 t 12 others
ORIGIN	
Query Match	3.5%; Score 144.4; DB 51; Length 597;
Best Local Similarity	67.8%; Pred. No. 2.5e-20;
Matches 202; Conservative	0; Mismatches 96; Indels 0; Gaps 0;
OY 3784	ttgtatggcaacagcggaagaatcgcgtagtcagctcgagattgtcttaccac 3843
Dd 9	TAGATTGGCTCAAAAGAAAAGAATTAACCCATGAGTGCGCATGTGGGCTGAGACTTC 68
OY 3844	atgatcatcgtcgatccgaagatcatcaccgtgactgaataaacgacaacacctctag 3903
Dd 69	ATGAACCTTGCATCCTTAGATCATTCATCGAGATGTAAAGCTGCAATGTATTATGG 128
OY 3904	acgaagaattcgaagcgattgttgagatttcggtgttgccaagtctaagactcataag 3963
Dd 129	ATGAAGATTGTGAAGCAGTTGTCGGACTTTGGCCTTGCCAAGTTGTGATGTAGGCG 188
OY 3964	aaactcacgtgacaacagcagtcgtgacacatcgatcacatcgcttcagaaatatct 4023
Dd 189	GGACAGATGTGACAACATCAAGTTCGGGGGCAACATGGGACACATATGACCACCGAGTACTTG 248
OY 4024	caaccggaataattctcagaagaacacgagatttcggtacagatcattgctctcaga 4081
Dd 249	CCACTGGGAAGTCATGAAAGGAGGATGTTTGTTATGGGATTAATGCTCTCTAGA 306
RESULT	15
LOCUS	A1729440
DEFINITION	BMHG13383 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U93048) somatic embryogenesis receptor-like kinase [Daucus

carota], mRNA sequence.  
 accession AI729440  
 version AI729440.1 GI:5048292  
 keywords EST.  
 source upland cotton.  
 organism *Gossypium hirsutum*  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eustosids II; Malvales; Malvaceae; *Gossypium*.  
 1 (bases 1 to 703)  
 Blewitt, M., Matz, E. C., Davy, D. F. and Burr, B.  
 ESTs from developing cotton fiber  
 Unpublished (1999)  
 On Jun 5, 1998 this sequence version replaced gi:318791.  
 contact: Ben Burr  
 Biology Department  
 Brookhaven National Laboratory  
 Upton, NY 11973, USA  
 Tel: 516-344-3396  
 Fax: 516-344-3407  
 Email: burr@bnl.bnl.gov  
 Seq primer: 73 primer.  
 Location/Qualifiers  
 1..703  
 /organism="Gossypium hirsutum"  
 /cultivar="Acala Maxxa"  
 /db\_xref="taxon:3635"  
 /clone\_lib="Six-day Cotton fiber"  
 /tissue\_type="Immature fiber"  
 /dev\_stage="Six days post anthesis"  
 /lab\_host="XLI-Blue"  
 /note="Vector: pBluescript II KS+"  
 188 a 148 c 180 g 186 t 1 others  
 BASE COUNT  
 ORIGIN

Query Match 3.5%; Score 144.4; DB 51; Length 703;  
 Best Local Similarity 67.8%; Pred. No. 2.5e-20;  
 Matches 202; Conservative 0; Mismatches 96; Indels 0; Gaps 0.

3784 ttgttgccgaacgcggaagagatcgcgcagtcgagtcgcgaagttgtcttactac 3843  
 |||||  
 9 TAGATTGGCTCTCAAAAGAAAAAGATTAACCCATGAGTGCAGCATGTGGGCTGGATACCTC 68  
 3844 atgatactcgtgcgcgaagatcaatcaccgtgacgttaaaacgcaaacatcctctag 3903  
 |||||  
 69 ATGAACATTCATCATCTTAAGATCATTCATTCAGATGGAAGGCTGCAAAAGTATTATTGG 128  
 3904 acggaagatccgaagcgggtgtgttgagatttcgggttgccaaactaatgtactataaag 3963  
 |||||  
 129 ATGAAGATTTTGAAGCAGTGTGGGTGAGTATTTGGCTTGCGAAATTTGTGATGTGAGGC 188  
 3964 acaatccggtgaacaacagcagtcgtygcacatcgtgcacatcgtctccgaatatctct 4023  
 |||||  
 189 GGACAGATGTGACACATCAAGTTCCGGGGGACAAATGGGACACATATGCAACCGAGTACTTG 248  
 4024 caaccggaataatctcaagsgaaacccgacgtttctcgatacggaatcatctctcaga 4081  
 |||||  
 Db 249 CCACTGGGAGAGTATCATAGAAAGAACGAGATCTTTGGTTATGGGATATATGCTCTTACA 306

Search completed: June 23, 2000, 19:06:25  
 Job time: 27418 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:12:15 ; Search time 12463.9 Seconds  
(without alignments)

-86.322 Million cell updates/sec

Title: US-09-180-798-22

Perfect score: 1106

Sequence: 1 tcgaccaccagctcgcgcga.....aaaaaaaaaaaaaaaaaaaaa 1106

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : GenEmbl:\*\*  
1: gb\_ba1:\*\*  
2: gb\_ba2:\*\*  
3: gb\_om:\*\*  
4: gb\_ov:\*\*  
5: gb\_pat:\*\*  
6: gb\_ph:\*\*  
7: gb\_pl1:\*\*  
8: gb\_pl2:\*\*  
9: gb\_pl3:\*\*  
10: gb\_pl4:\*\*  
11: gb\_pl5:\*\*  
12: gb\_pl6:\*\*  
13: gb\_pl7:\*\*  
14: gb\_pl8:\*\*  
15: gb\_pl9:\*\*  
16: gb\_pl10:\*\*  
17: gb\_pl11:\*\*  
18: gb\_pl12:\*\*  
19: gb\_pl13:\*\*  
20: gb\_pl14:\*\*  
21: gb\_pl15:\*\*  
22: gb\_pl16:\*\*  
23: gb\_pl17:\*\*  
24: gb\_pl18:\*\*  
25: gb\_pl19:\*\*  
26: gb\_pl20:\*\*  
27: gb\_pl21:\*\*  
28: gb\_pl22:\*\*  
29: gb\_pl23:\*\*  
30: gb\_pl24:\*\*  
31: gb\_pl25:\*\*  
32: gb\_pl26:\*\*  
33: gb\_pl27:\*\*  
34: gb\_pl28:\*\*  
35: gb\_pl29:\*\*  
36: gb\_pl30:\*\*  
37: gb\_pl31:\*\*  
38: gb\_pl32:\*\*  
39: gb\_pl33:\*\*  
40: gb\_pl34:\*\*  
41: gb\_pl35:\*\*  
42: gb\_pl36:\*\*  
43: gb\_pl37:\*\*  
44: gb\_pl38:\*\*

45: gb\_pl39:\*\*  
46: gb\_pl40:\*\*  
47: gb\_pl41:\*\*  
48: gb\_pl42:\*\*  
49: gb\_pl43:\*\*  
50: gb\_pl44:\*\*  
51: gb\_pl45:\*\*  
52: gb\_pl46:\*\*  
53: gb\_pl47:\*\*  
54: gb\_pl48:\*\*  
55: gb\_pl49:\*\*  
56: gb\_pl50:\*\*  
57: gb\_pl51:\*\*  
58: gb\_pl52:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1106	100.0	1106	5	A67817	A67817 Sequence 22
2	924	83.5	981	5	A67819	A67819 Sequence 24
3	912.4	82.5	1063	5	A67825	A67825 Sequence 30
4	714.2	64.6	789	5	A67821	A67821 Sequence 26
5	330	29.8	894	5	A67823	A67823 Sequence 28
6	191	17.3	2089	5	A67827	A67827 Sequence 32
7	151.2	13.7	936	8	SB062279	U62279 Sorghum b1c
8	96.6	8.7	4604	7	LELPGENE	X95269 L.esculentu
9	84.2	7.6	1755	8	DCU93048	U93048 Daucus caro
10	79.8	7.2	8153	50	A67797	A67797 Sequence 2
11	77.8	7.0	8440	7	ABO10075	AC021198 Arabidops
12	77.8	7.0	118086	7	ATF6H1	ABO10075 Arabidops
13	76.6	6.9	5033	8	INU77888	AL021684 Arabidops
14	75.8	6.9	142418	7	AP000815	U77888 Ipomoea nil
15	74	6.7	6695	5	A67796	AP000815 Oryza sat
16	73.6	6.7	101284	8	ATNAC009991	A67796 Sequence 1
17	72.4	6.5	4081	5	A67815	AC009991 Arabidops
18	72.4	6.5	98471	8	AC012654	A67815 Sequence 20
19	71.8	6.5	100887	50	ATAC011620	AC012654 Arabidops
20	70.8	6.4	3176	5	EL12705	AC011620 Arabidops
21	70.8	6.4	3176	8	ATU47029	EL12705 Arabidops
22	70.4	6.4	112529	44	AC015446	U47029 Arabidops
23	69.6	6.3	116944	8	ATAC011765	AC015446 Arabidops
24	69.4	6.3	116944	8	AC008017	AC011765 Arabidops
25	67.4	6.1	2786	7	ABO29327	AC008017 Arabidops
26	67.2	6.1	88401	8	F23M19	ABO29327 Nicotiana
27	66.8	6.0	69142	7	ABO12239	AC007454 Arabidops
28	66.4	6.0	96475	7	ATF17M5	ABO12239 Arabidops
29	66.4	6.0	111945	8	ATF1C12	AL035678 Arabidops
30	65.6	5.9	82033	7	ABO11476	AL022224 Arabidops
31	65.6	5.9	86139	7	ABO28616	ABO11476 Arabidops
32	65.2	5.9	2278	8	SBRLK1	ABO28616 Arabidops
33	64.8	5.9	108355	50	AC005957	Y14600 Sorghum b1c
34	64.4	5.8	61712	7	ABO17061	AC005957 Arabidops
35	64.4	5.8	133840	8	AC003981	ABO17061 Arabidops
36	64.2	5.8	3573	5	A57133	AC003981 Complete
37	64	5.8	3368	8	U42445	A57133 Sequence 4
38	64	5.8	6471	5	A57130	U42445 Lycopersico
39	64	5.8	6471	5	U42444	A57130 Sequence 1
40	63.4	5.7	87434	7	ABO10698	U42444 Lycopersico
41	63.4	5.7	23143	8	F15P11	ABO10698 Arabidops
42	62.6	5.6	65899	50	AC006436	AF160760 Arabidops
43	62	5.6	101154	8	ATF1P2	AC006436 Arabidops
44	61.6	5.6	3979	5	A67428	AL132935 Arabidops
45	61.2	5.5	3979	5	A67428	A67428 Sequence 1

## ALIGNMENTS



```

RESULT 1
LOCUS A67817 1106 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 22 from Patent WO9743427.
ACCESSION A67817
VERSION A67817.1 GI:4756639
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1106)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICETIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source
location/Qualifiers
1..1106
/organism="unidentified"
/db_xref="taxon:32644"
142..798
/note="unnamed protein product"
CDS
/protein_id="CAB42244.1"
/db_xref="GI:4756640"
/translation="MASRNYRWELFASLTLTLALILHVEANSEGDALYALRRSLTDP
DHVQSDPTLVNPTCFHYTCNDNRVTVDLSNSGLAPELGKLEHLYELY
KNNIGTIPSELGNLKNLISDLNNNTGIVPTSLGKLSLVFLRLNDRLTGPILP
HSRQSQAFKVVVSSNDLGTIPNGPFAHPILOWENNPRLEGPPELLGLASYDINCT
"
BASE COUNT 331 a 258 c 206 g 311 t
ORIGIN

```

```

Query Match 100.0%; Score 1106; DB 5; Length 1106;
Best Local Similarity 100.0%; Pred. No. 5.8e-255;
Matches 1106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 tggaccacgagtcgctcaactcaataaagggaacacgtaaccctaattgct 60
DB 1 tggaccacgagtcgctcaactcaataaagggaacacgtaaccctaattgct 60
QY 61 ttctccctctgttcaagaanaatttccttactctcaaatccctttccctct 120
DB 61 ttctccctctgttcaagaanaatttccttactctcaaatccctttccctct 120
QY 62 ttctccctctgttcaagaanaatttccttactctcaaatccctttccctct 120
DB 62 ttctccctctgttcaagaanaatttccttactctcaaatccctttccctct 120
QY 121 cttaaacctcgaagctcacatgagctcgcgaacatacgttggagctcttcgacgt 180
DB 121 cttaaacctcgaagctcacatgagctcgcgaacatacgttggagctcttcgacgt 180
QY 121 cttaaacctcgaagctcacatgagctcgcgaacatacgttggagctcttcgacgt 180
DB 121 cttaaacctcgaagctcacatgagctcgcgaacatacgttggagctcttcgacgt 180
QY 181 tegttaaaccttaacctagcttgaattcaacctgctcgaagcaaacctcgaaggaatgct 240
DB 181 tegttaaaccttaacctagcttgaattcaacctgctcgaagcaaacctcgaaggaatgct 240
QY 181 tegttaaaccttaacctagcttgaattcaacctgctcgaagcaaacctcgaaggaatgct 240
DB 181 tegttaaaccttaacctagcttgaattcaacctgctcgaagcaaacctcgaaggaatgct 240
QY 241 ctctaagctcttcgcgcgaagtttgcacagatccagacacatgctcctcgaagctggatcca 300
DB 241 ctctaagctcttcgcgcgaagtttgcacagatccagacacatgctcctcgaagctggatcca 300
QY 241 ctctaagctcttcgcgcgaagtttgcacagatccagacacatgctcctcgaagctggatcca 300
DB 241 ctctaagctcttcgcgcgaagtttgcacagatccagacacatgctcctcgaagctggatcca 300
QY 301 actctgttgaatccctgtacccgtgttccatgctcaccgttaaccaaagaacacgctgact 360
DB 301 actctgttgaatccctgtacccgtgttccatgctcaccgttaaccaaagaacacgctgact 360
QY 301 actctgttgaatccctgtacccgtgttccatgctcaccgttaaccaaagaacacgctgact 360
DB 301 actctgttgaatccctgtacccgtgttccatgctcaccgttaaccaaagaacacgctgact 360
QY 361 cgtgtgagatttgggaatccaacctcctcgtgacaatcttgcgcctggagcttgggaagctt 420
DB 361 cgtgtgagatttgggaatccaacctcctcgtgacaatcttgcgcctggagcttgggaagctt 420
QY 361 cgtgtgagatttgggaatccaacctcctcgtgacaatcttgcgcctggagcttgggaagctt 420
DB 361 cgtgtgagatttgggaatccaacctcctcgtgacaatcttgcgcctggagcttgggaagctt 420
QY 421 gaacattacaglatcagagctctacaaaacacacatccaagaatactaccttcgaa 480
DB 421 gaacattacaglatcagagctctacaaaacacacatccaagaatactaccttcgaa 480
QY 421 gaacattacaglatcagagctctacaaaacacacatccaagaatactaccttcgaa 480
DB 421 gaacattacaglatcagagctctacaaaacacacatccaagaatactaccttcgaa 480
QY 481 cttggaatactgaagaatctcagctgagctgtgtatacaacaacatcttcgaagata 540
DB 481 cttggaatactgaagaatctcagctgagctgtgtatacaacaacatcttcgaagata 540
QY 481 cttggaatactgaagaatctcagctgagctgtgtatacaacaacatcttcgaagata 540
DB 481 cttggaatactgaagaatctcagctgagctgtgtatacaacaacatcttcgaagata 540

```

```

QY 541 gtcccaacttcttgggaanaattgaagctcctgctcttcttaacgcttaacagacacga 600
DB 541 gtcccaacttcttgggaanaattgaagctcctgctcttcttaacgcttaacagacacga 600
QY 601 ttgaacgctcacaatcctcagaagcactcagcgaatcccaagccttaagaattgttgaagctc 660
DB 601 ttgaacgctcacaatcctcagaagcactcagcgaatcccaagccttaagaattgttgaagctc 660
QY 601 ttgaacgctcacaatcctcagaagcactcagcgaatcccaagccttaagaattgttgaagctc 660
DB 601 ttgaacgctcacaatcctcagaagcactcagcgaatcccaagccttaagaattgttgaagctc 660
QY 661 tcaagcaatgatcttgttgggacaatcccaacaacaggaaccccttggccaatccctta 720
DB 661 tcaagcaatgatcttgttgggacaatcccaacaacaggaaccccttggccaatccctta 720
QY 661 tcaagcaatgatcttgttgggacaatcccaacaacaggaaccccttggccaatccctta 720
DB 661 tcaagcaatgatcttgttgggacaatcccaacaacaggaaccccttggccaatccctta 720
QY 721 cagaacttgggaacaccccgagatttggagagcagcgaattactcgtgcttcgaagctac 780
DB 721 cagaacttgggaacaccccgagatttggagagcagcgaattactcgtgcttcgaagctac 780
QY 721 cagaacttgggaacaccccgagatttggagagcagcgaattactcgtgcttcgaagctac 780
DB 721 cagaacttgggaacaccccgagatttggagagcagcgaattactcgtgcttcgaagctac 780
QY 781 gaactaactgacactgaaacacactgcaaaaacactgaaatgaagaattggggggtagacc 840
DB 781 gaactaactgacactgaaacacactgcaaaaacactgaaatgaagaattggggggtagacc 840
QY 781 gaactaactgacactgaaacacactgcaaaaacactgaaatgaagaattggggggtagacc 840
DB 781 gaactaactgacactgaaacacactgcaaaaacactgaaatgaagaattggggggtagacc 840
QY 841 ttgtaagaacacttcaacacttataatcaatcattatgtatgaatgaatata 900
DB 841 ttgtaagaacacttcaacacttataatcaatcattatgtatgaatgaatata 900
QY 841 ttgtaagaacacttcaacacttataatcaatcattatgtatgaatgaatata 900
DB 841 ttgtaagaacacttcaacacttataatcaatcattatgtatgaatgaatata 900
QY 901 ttragaataaacaacaaatgaagaatcgtaacatcgtaacatcgtaacatcgtaacatcgtaac 960
DB 901 ttragaataaacaacaaatgaagaatcgtaacatcgtaacatcgtaacatcgtaacatcgtaac 960
QY 901 ttragaataaacaacaaatgaagaatcgtaacatcgtaacatcgtaacatcgtaacatcgtaac 960
DB 901 ttragaataaacaacaaatgaagaatcgtaacatcgtaacatcgtaacatcgtaacatcgtaac 960
QY 961 aacttcgagctcgtatgtaaaattcctaactgagatttcgctactgtaactgctgct 1020
DB 961 aacttcgagctcgtatgtaaaattcctaactgagatttcgctactgtaactgctgct 1020
QY 961 aacttcgagctcgtatgtaaaattcctaactgagatttcgctactgtaactgctgct 1020
DB 961 aacttcgagctcgtatgtaaaattcctaactgagatttcgctactgtaactgctgct 1020
QY 1021 tttgggattcttgggaagtaacatttgaattgtaagtatacagtttgccttgccttgcct 1080
DB 1021 tttgggattcttgggaagtaacatttgaattgtaagtatacagtttgccttgccttgcct 1080
QY 1021 tttgggattcttgggaagtaacatttgaattgtaagtatacagtttgccttgccttgcct 1080
DB 1021 tttgggattcttgggaagtaacatttgaattgtaagtatacagtttgccttgccttgcct 1080
QY 1081 gcaaaaaaaaaaaaaaaaaaaaaa 1106
DB 1081 gcaaaaaaaaaaaaaaaaaaaaaa 1106

```

```

RESULT 2
LOCUS A67819 981 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 24 from Patent WO9743427.
ACCESSION A67819
VERSION A67819.1 GI:4756641
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 981)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICETIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source
location/Qualifiers
1..981
/organism="unidentified"
/db_xref="taxon:32644"
104..760
/note="unnamed protein product"
CDS
/protein_id="CAB42246.1"
/db_xref="GI:4756642"
/translation="MASRNYRWELFASLTLTLALILHVEANSEGDALYALRRSLTDP
DHVQSDPTLVNPTCFHYTCNDNRVTVDLSNSGLAPELGKLEHLYELY
KNNIGTIPSELGNLKNLISDLNNNTGIVPTSLGKLSLVFLRLNDRLTGPILP
ALTPSLKVVVSSNDLGTIPNGPFAHPILOWENNPRLEGPPELLGLASYDINCT
"
BASE COUNT 286 a 236 c 180 g 279 t
ORIGIN

```







Db 483 CAACAAACATCTTACAGGAGTAGTCCACTTCTTGGGAAATGAAAGTCTGCTCT 542  
Qy 579 ttaacggctaataagcaacgattgacggctccaat-cctaagagcctaagcgcaatcc 637  
Db 543 TTTAGGGCTTAATGACCAACCGATTGACGGGCCAATCCCTAGACACTGCAATGCC 602  
Qy 638 aagcccttaagttgttgcacgtctcaagcaatgattgtgttggaacatccacaaga 697  
Db 603 AAGGC-TTAAAGTGTGATGCTCAAGCAATGATTGTGTGGAACAATCCCAACAAAG 661  
Qy 698 gaeccttgctcacatctccttcaagaacttgaagaacacccagattggaggacgg 757  
Db 662 GACCTTGTCTCAGATCCCTTACAGAACTTTGAGAAACCCGAGGTTGAGAGGCCGG 721  
Qy 758 aattactcgtcttgcaagtagaacaactacgcacactgaacaacttgcgaactga 817  
Db 722 AATTACTGGGCTTGGCAAGCTACGACACTACGACCTGAAAAAATTTGGCAAACTGA 781  
Qy 818 aaatgaagaattgggggtgacctgtgaagaacacattcacacttatacaatcacat 877  
Db 782 AAATGAAGATTTGGGGGCTACCTGTAAAGAACATTCACACTTATCAATATACAT 841  
Qy 878 ctattatgataatagatataatgtgaagaaaaaagaagaagaatcgatcgat 937  
Db 842 CTACTATGATATATATATATATGATGTCAA-AAAAAATGAAAGATCGAATCAGT 899  
Qy 938 aatactacgtcgtcaattgagaacttgcaggtc--tgaatgaatttcaattcgaatg 995  
Db 900 AATATCATCTGGTCTCATTTGAGAACTTTGAGTCTGTATTAATTTCTAAATGGG 959  
Qy 996 atttcgctactgtaattgtctgttggaattctgaagaagaacattgtatgtat 1055  
Db 960 ACTTCCGCTACGATATGTTGCGTGTGGGATTCGAGAGATGACATTTGATTTGAT 1019  
Qy 1056 ggtatcaagtgtctgcctgtctgcgaaaaaaataaaaaa 1099  
Db 1020 GGTATCAAGTTGTCTGCTGTGCAAAAAAATAAAAAA 1063

RESULT 4  
A67821 789 bp DNA PAT 05-MAY-1999  
LOCUS A67821  
DEFINITION Sequence 26 from Patent WO9743427.  
ACCESSION A67821  
VERSION A67821.1 GI:4756643  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 789)  
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE PRODUCTION OF APOMICITIC SEED  
JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
source  
1. 789  
Location/Qualifiers  
/organism="unidentified"  
/db\_xref="taxon:32644"  
CDS  
<2..664  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42248.1"  
/db\_xref="GI:4756644"  
/translation="RPTRRNRYRWELFASLLITLALTHVANESEDAATYALRSLL  
DPDRHQSMDPTLVNDCPTFFHYTCANDNRYTVYDGNLSLSEHLPELGRLEHLYLE  
LTKNNIGTIPSELGNLKNLISLDLYNNNTLIVPSLGRKLSVLERLNRLTGPI  
PRALTAIPSLKAVDVSSNDLGGTIPNGFAHPIPLQNEENNRLBGPILLGLASDYDN  
CT"

BASE COUNT 235 a 191 c 156 g 207 t  
ORIGIN  
Query Match 64.6%; Score 714.2; DB 5; Length 789;

Best Local Similarity 96.2%; Pred. No. 4,2e-161;  
Matches 753; Conservative 0; Mismatches 28; Indels 2; Gaps 2;  
Qy 139 cacatggcgtctgaaactatcgttggagctcttcagacttcgttaaccctaacta 198  
Db 5 CCCAGGCTCCGCAAACTATCGGTGGAGCTCTCCACCTTCGTTAAATCTAACCTTA 64  
Qy 199 gcttggatcacctggtcgaaggaactccgaaggaatgctctcagctcttcggcg 258  
Db 65 GCTTGGATTCACCTGTGGGAAAGCAACTCCGAAAGAAATGCTTTAGGCTCTTGGCGG 124  
Qy 259 agttgacagatccagaccatgctctcagagctggagatccaaactctgttaactctgt 318  
Db 125 AGTTTAAAGATCCGAGACATGTTCTCAGAGCTGGGATCCAACTCTGTTAATCCTGT 184  
Qy 319 accgtgttcacatgtaacctgttaacaaacacccgctacactcgtgtgttggaat 378  
Db 185 ACCTGCTCCATGTCACCTGTAAACCAAGAACCGGCTCACTGCTGTGATTTGGGGA 244  
Qy 379 tcaaacctctctgacatcttgagcctgagctgagcttggagattgaacttaacatcta 438  
Db 245 TCAACCTCTCTGGACATCTTGCGCTGAGCTTGGAAAGCTTGAACTTACAGTATCTA 304  
Qy 439 gagctctcaaaaacaacatccaaaggaactataccttcgaacttggaaacttgaagaat 498  
Db 305 GAGCTCTACAAAACAACATCCAAAGAACTATACCTTCGGAAGTGGAAATCTGAAGAA 364  
Qy 499 ctcaatcagcttggatcgttaacaacaactctcaaggagatgtccacttctcggga 558  
Db 365 CTCAATCACCTTGGATCTGTACAAACAACTTACAGGAGTATGTTCCCACTTTCTTGGGA 424  
Qy 559 aaatgaagctctcgtgtcttttcttaacgcttaagaaacacgattgagcgttccaat-cct 617  
Db 425 AATTTGAAGTCTCTGCTTTTACGGCTTATACAAACGATTTGACGGGCCAATCCT 484  
Qy 618 agagactcagcgaactcccaagcctttaagttgttgagctctcgaagcaatgatgtgt 677  
Db 485 AGAGACATCATGCATGCCAAGCC-TTAAAGTGTGATGTGTCAAGCAATGATTTGTG 543  
Qy 678 tggagcaatcccaacaacaggaaccttgcacatccctttacaagaacttgaagaaca 737  
Db 544 TGAACAAATCCCAACCAACGGAACCTTTGCTCACATTCCTTACAGAACTTGAAGAAC 603  
Qy 738 cccgagattgaggagaccggaattactcgtgtcgaagtagaagaactaactgcaactg 797  
Db 604 CCCGAGTTGGAGGACCGGAATTAAGTCTTGGCAAGCTACACACTACGACACTG 663  
Qy 798 aaacacgtgcaaaactgtaaatgaagaattgggggtgacacttgaagaacactcac 857  
Db 664 AAAAAATTGGCAAAACCTGAATAATGAAGAAATTTGGGGGTGACCTGTAAAGAACACTTCAC 723  
Qy 858 caacttacaatatcacatctatgtgaataagatataatgtgtgaataaacaacaaa 917  
Db 724 CACTTATCAAAATATCAATCTACTATATATATATATATATATATATATATATATATAT 783  
Qy 918 aaa 920  
Db 784 AAA 786

RESULT 5  
A67823 894 bp DNA PAT 05-MAY-1999  
LOCUS A67823  
DEFINITION Sequence 28 from Patent WO9743427.  
ACCESSION A67823  
VERSION A67823.1 GI:4756645  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 894)  
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE PRODUCTION OF APOMICITIC SEED



```

JOURNAL      Patent: WO 9743427-A-20-NOV-1997/
              CIBA GEIGY AG (CH)
FEATURES
  source      location/qualifiers
              1..894
              /organism="unidentified
              /db_xref="taxon:3264"
              <1..678
CDS

```

BASE COUNT	270 a	163 c	176 g	28
ORIGIN				

Query Match	29.8%;	Score 330;	DB 5;	Length 894;
Best Local Similarity	71.9%;	Pred. No. 4.5e-69;		
Matches 532;	Conservative	0;	Mismatches 115;	Indels 93;
				Gaps 4;

QY	225	ctccgaagggaagcctcttaagcctcttcgcggaggttggcaatccaaacctgcct	284
Db	15	CTCCGAAGGGGAGCGCTTTCACGCGCTTCCGGAGCTTATCAATCCAGACAATGTTG	74
QY	285	ccagagctggatccaaactcttgtaatcccttgtaacctggtgtccatgtaacctgtaacca	344
Db	75	TCAGAGTTGGGATCCCAACTCTGTATATCTCTGTACTGTGTTTCATGCTCACTGTATCA	134
QY	345	agacaacgggtacccgctgtagtttggaaattcacaacctccctgaacctctggcc	404
Db	135	ACACCAATCAGTACATCTGCTGTGGATTGGGGAATTCAACTTATCTGACATCTAGTACC	194
QY	405	tgaagcttgggaagcttgaacaattacagtatct-----	438
Db	195	TGAAGCTGGGAACCTTGACAATTATCAATCTGTATGATCAATCACTCTTTGCTTT	254
QY	438	-----	438
Db	255	TGATTATCTAAAAACATTTACATTATCACTCAACATTTAAACATTTTGCTTTGAGTCATA	314
QY	438	---agagcttacaanaaacacacaccacaagaagactaatacttcggaacttggaaattgaa	494
Db	315	TAGTGAAGCTCTACAAAAACGATGATCAAGAACCTATACCTCTGTGAGCTGGAAATCTGAA	374
QY	495	gaactcatatgaacttggatctgtacacaacaactttacaggatagttcccaattcct	554
Db	375	GAGCTATATAGTTGGATCTGTACAAACAACATCTCCACGGGAAAAATCCCATTTCTTT	434
QY	555	ggggaaaattgaagctctcgtgctctttaaagcttaataagacaacgcatggacggtcc-aa	613
Db	435	GGGAAATGTGAAGTCACTGTTTTTTTTGGGCTTTAAGCAAAAACGATTGACGGTCTCAT	494
QY	614	tccctaggacactcaagggcaatcccaagcctttaaagttgttgaagtcacgaacatgatt	673
Db	495	TCCTAGGGAATCTCAACATTTATTTCAAGGCC-TTTAAAGTTGTTGATCTCTCAAGGATATATT	553
QY	674	tgtgtgggaacatcccaacaacggaaccccttgcctcaactcctttaaagaacttga	733
Db	554	TGTTGGAACATTTCCAGTGTGAAGAGACCTTTTGAACACATTCCTATGCAAAACTTTGAGA	613
QY	734	acaaaccggaattggagggacggaattactcgtgtcttgcgaagctaagaaactaactga	793
Db	614	ACAACCTGGAAGTTGGAGGGACCAAGACTCTAGGCTTTGCGAGCTATACACCAATTGCA	673
QY	794	ccctgaacaactggcaaaccttgaaaatgaagaattgggggttgaaccttgaagaacct	853
Db	674	CTTAAAAAGAGTTGAAGAACCTATTAAGAAGAA-TGTTAGTGACCTTTGAABAACCTCT	732
QY	854	tcaccactttatacaaatatc	873

Db 733 GTACCAAGTGTGTAATC 752

RESULT	6		PAT	05-MAY-1999
A67827		A67827	2089 bp	DNA.
LOCUS		Sequence 32 from Patent WO97343427.		
DEFINITION		A67827		
ACCESSION		A67827.1	GI:4756649	
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
		thale cress.		
		Arabidopsis thaliana		

## REFERENCE 1 (bases 1 to 2089)

**AUTHORS** De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.

JOURNAL  
Patent: WO 9/43427-A 20-NOV-1997;  
CIPA CEECV AC (CN)

## FEATURES

### Location/Qualifiers

source 1. .2089

```
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="SERK GENE CDNA"
```

CDS

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42254.1"
/db_xref="GI:4756650"
/translation="MEESSVYVETLLISDILPNHSMILASANEGLATKTAFTAVDYDNR
AVLQGRDPTLVNPGCTHFVYTCNNEDSYTRIDGNALMSCHLVPELQKNIQYLAETLS
NNITQIPISNIGNLTNLVSUQLTLNLSGGLPESIGLSKLFRLNNLNSLTGSIIPWS
LITITLQVLDSNNRSLGSGVPNGSFSLTPEAFANIMDLGCTPSHSCGSPSEPP
PPPELPPPVTPSPSGYITGAIGAVGAAGALFAFPAALFMMWRSPDLIFPDVPA
EDDPVHLGQKREFSLRELVASDSGFSNNKILLDGEGFVGYGRGLADTLVAVNRKEK
EETPGCELFQTEVEEMISMAVHNRLLRLRFQCTPERLVLYPYANSTGASVCRERK
EATMDQWPTKRRRIALGARGSLVLDHDOPTKILHDYQAAIIMLDEFEVQAVDGL
AKIMQKTHYTTAVRGSTIGHIAPVLSYKDSSEKRDYGCITIMLEITVQRAPFDA
RLANDDDMLDMWYKGLKEKRLBMLDPLQTNIBERLLEQVTVALLCTQSGPMEE
PMSSSVYVETLLISDILSPNPSDWIMLIDSTYNLHAIVELSGSRR"

```

Query Match	17.38;	Score 191;	DB 5;	Length 2089;
Best Local Similarity	60.5%;	Pred. No. 9e-36;		
Matches 314;	Conservative	0;	Mismatches 205;	Indels 0;
			Gaps	0;

OY	220	gcaaacctccgaagaagatcgtctctcaagctctctccgcggagtttgaaagatccaagccat	279
Db	270	gctatatttgamaagtgatgctttgcatcttggacgttgaaagggttactctagttgattccaaacaaat	329
OY	280	gtccctccagaagctcggagatccaaactctgttaacctcttaccttggttccatgtccacatgt	339
Db	330	gctttgcagagcgtggagatccctacgcagtaggaatcctttcacatgattgctcatgtccactgtgc	389
OY	340	aacaaagacacacgcgcctactcgtctgtgattttgggaatttcaaacccctctggaacatct	399
Db	390	aacacagagaaacgtgtctatnagaagttgattttgggggaatgcagagtttatctggccatttta	449
OY	400	gcgcgttaagtttgggaagcttgaacatttaacgtatctctagagctctcaaaaacaacatc	459
Db	450	gttccaaagccttgcgtgctcagaaattttgcagattttggagcttttacaattaacacata	509
OY	460	caaggaactataccctccgaacttggaaactctgaagaactctcatcagcttgatctgtac	519
Db	510	actgagcccgattcctctgataattcttggaataatgcacaaactttagagattttggactttttac	569
OY	520	aacaacaattcttcaggaagtagtcccaactcttcttcgggaanaattgaagctctcgtgcttt	579



Db	570	TTAAACGCTTCCCGTCTCTATTCGGAGATGATGGAAAGCCTTCAAGCTTACAGTTT	629
Qy	580	ttacggtcttaagcaacacogattgacacggttccaatccattagacgtcaccgaatcccaa	639
Db	630	CTCCGGCTTAAACAACAAGTGTCACTGGGTCAATTCCTATGTCACGTACCAATATTACT	689
Qy	640	gccttaaatgtttgaagtcacgaacatgttggttgagacatcccaaacagga	699
Db	690	ACCTTAAAGTGTAGTGTATCATCAATTAACAACACTCTCGTTCAGTTCTCGAATGCG	749
Qy	700	cccttgctacattcccttacagaaacttgagaaac	738
Db	750	TCCTTCACTTTCACACCAACATGATTGGCTAATAAC	788
RESULT	7		
SBUS62279	936	bp	PLN
LOCUS	SBUS62279	936	bp
DEFINITION	Sorghum bicolor leucine-rich repeat-containing extracellular		
ACCESSION	U62279		
VERSION	U62279.1		
KEYWORDS	GI:1710123		
SOURCE			
ORGANISM	Sorghum		
REFERENCE	Sorghum bicolor		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;		
JOURNAL	Poaceae; Sorghum		
REMARK	1 (bases 1 to 936)		
REFERENCE	Hipskind, J.D., Nicholson, R.L. and Goldsbrough, P.B.		
AUTHORS	Isolation of a cDNA encoding a novel leucine-rich repeat motif from		
TITLE	Sorghum bicolor inoculated with fungi		
JOURNAL	Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)		
REMARK	97124217		
REFERENCE	Erratum: [published erratum appears in Mol Plant Microbe Interact		
AUTHORS	1997 Mar;10(2):302]]		
TITLE	2 (bases 1 to 936)		
JOURNAL	Hipskind, J.D.		
REMARK	Direct Submission		
TITLE	Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant		
JOURNAL	Pathology, Purdue University, West Lafayette, IN 47907, USA		
FEATURES			
SOURCE	location/Qualifiers		
CDS	1..936		
	/organism="Sorghum bicolor"		
	/db_xref="taxon:4558"		
	/clone="SLRR"		
	12..755		
	/function="protein-protein ligand binding"		
	/note="leucine-rich repeat-containing extracellular		
	glycoprotein; contains six N-glycosylation sites		
	[NX(S/T)]		
	/codon_start=1		
	/protein_id="AAC49559.1"		
	/db_xref="GI:1710124"		
	/translation="NAPQFAAGFLTGLALATFASCTEGEDILYKQRLAMEDPNVNL		
	OSGNSNLANPCMEHYTCNNNFVIRVDLAGNAGISGLPDLAIOIYELKXGL		
	NGSIPETLGNLTNLTISIDMLNLITGEIPTTISITRYLRLLQNNDLTPITSSGN		
	LTSLSKRLQENLSGALPASLGNIKALQFSLNDMLTGTVPKSPSTFQNTL		
	NTRNNLIDGTRTSSSGLRVATIIQDALKTA"		
	36..77		
	/evidence=not_experimental		
BASE COUNT	247 a	232 c	207 g
ORIGIN	250 t		
Query Match	13.7%	Score 151.2;	DB 8; Length 936;
Best Local Similarity	61.4%	Pred. No. 3.1e-26;	
Matches 243; Conservative 0; Mismatches 153; Indels 0; Gaps 0;			
Qy	223	aactcgaagagatgctctcttaagctcttcgccgagattgacagatccacacatgctc	282
Db	81	AACACTGAAGGTGACATACGTGTAACGCAAGGTTGGCATGGAGAGACCAACAACGTG	140



	NCS"	
intron	346..2380	
	/gene="LRP"	
	/number=1	
exon	2381..2452	
	/gene="LRP"	
	/number=2	
intron	2453..2564	
	/gene="LRP"	
	/number=2	
exon	2565..2708	
	/gene="LRP"	
	/number=3	
intron	2709..2927	
	/gene="LRP"	
	/number=3	
exon	2928..2999	
	/gene="LRP"	
	/number=4	
intron	3000..3080	
	/gene="LRP"	
	/number=4	
exon	3081..3152	
	/gene="LRP"	
	/number=5	
intron	3153..4282	
	/gene="LRP"	
	/number=5	
exon	4283..>4604	
	/number=6	
BASE COUNT	1311 a 758 c 888 g 1647 t	
ORIGIN		

Query Match	8.7%	Score 96.6	DB 7	Length 4604
Best Local Similarity	69.1%	Pred. No. 3.7e-13		
Matches 132	Conservative 0	Mismatches 59	Indels 0	Gaps 0
QY 178	gtctcgtaaccctaacccttagctttagatctaacctggtcgaagaacaactccgaagagat	237		
Db 158	GGTTTGTGGCACTTCTTCTGTACTGTGGCTGTAAAGGGGAATTCAGAAAGGGGAT	217		
QY 238	gtctctcaagcgtcttcgcgcgagattgacagatccagaaccatgcttcctcagaagcttggat	297		
Db 218	GCTTTTGAGCCCTTGCGCGAGCTTACTGACCCGGGTAAACGGTTACAGAGCTGGAT	277		
QY 298	ccaactctgttaactctctgtgtaacctggtgttcatagttcaacctgttaaccaagaacaacgcgtc	357		
Db 278	CCAAATCTTTTAAACCTTGTACTGCTGCTTTCATGTCACTTGCAACGAGATATCAAGTT	337		
QY 358	actcgtgtgga 368			
Db 338	ACTGCTGTGA 348			
RESULT 9				
LOCUS DCU93048	1755 bp	mRNA	PLN	28-JUN-1997
DEFINITION	Daucus carota somatic embryogenesis receptor-like kinase mRNA,			
ACCESSION	U93048.1			
VERSION	GI:2224910			
KEYWORDS				
SOURCE	carrot.			
ORGANISM	Daucus carota			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
TITLE	eudicots; Asteridae: euasterids II; Apiales; Apiaceae; Daucus.			
JOURNAL	Schmidt E.D., Guzzo F., Toonen M.A. and de Vries S.C.			
	A leucine-rich repeat containing receptor-like kinase marks somatic			
	plant cells competent to form embryos			
	Development 124 (10), 2049-2062 (1997)			

BASE COUNT	ORIGIN	FEATURES	source	REFERENCE	MEDLINE
506 a	347 c	407 g	495 t	97313247	2 (bases 1 to 1755)
					Schmidl,E.D.L., Guzzo,F., Toonen,M.A.T. and de Vries,S.C.
					Direct Submission
					Submitted (12-Mar-1997) Molecular Biology, Agricultural University
					of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
					Location/Qualifiers
					1..1755
					/organism="Daucus carota"
					/db_xref="taxon:4039"
					34..1695
					/note="SERK"
					/codon_start=1
					/product="somatic embryogenesis receptor -like kinase"
					/protein_id="AAB61708.1"
					/db_xref="gi:2224911"
					/translation="MNRNINILNMYOFDAVIDIKRGVIMLTLEYLNNISGSPISPSIDG
					NITNIVASDILYMNFSGSPIDPTIGKRLRLRLNNNSISGPIPSMLNITITLVDLID
					SNNRISGVPDNGSFLFTPISTFRNNINLGGVPTGPRGSGSPSPPPPTIPSTVDP
					RSQNEELAVADPTSTILGRGGEKXKGLADGSLVAVKRLBERTGDEGPEVHLGOLK
					PLSPNLSPTAAGVAGAAALFLFAPAAAPARMMWRRRPRHFDVAEDPEVHLGOLK
					ALGISAIVHRNLLRLRGCPMTITRLLAYIPMANGSAVSLRRQSGSEPLDMPKRTKI
					AVGSLRGSLYHDCDPKRIHRDYKAANILLDEEFAVVGDPGLARLADYKDTPTTKA
					VGSLGTIAPLEYLSTGKSEKTEDEGVGIMLLELLIGQAPFLARLADDDVIMLDWV
					KLILKERTLEMLNDPDLNNYIDPEVQGLQVLLCTQSSPKERKMEYVRMLDGCG
					LAEKDEMQKVEYVTHQDEYLAAPHRTSWMILSTDINHAEELSGPR"

[illegible]



REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 1815)	De V.S., Schmidt, E.D., Van, H. G. and Hecht, V. F.	PRODUCTION OF APOMITIC SEED	PATENT: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)			
FEATURES			
source			
location/Qualifiers			
1..1815			
/organism="Daucus carota"			
/db_xref="taxon:4039"			
94...1755			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CABA2241.1"			
/db_xref="GI:4756624"			
/translation="MNRNSINILNTYMOFIDATYIDRKGYMTLELYSNNSISGRPSIDSG NLTNIVSLDLTNSFSFSPETPDITLKRILNRPFLNNNSLSGFIPLSLNITITLOVLD SNNRISGVPDNGSPSLFTPISEFANNILNLEGPVTPGRCSPSPFPPEFTIPSTVODP PQNDPTGLAIGVAAGAAALFALFAPAMAFWMRRRRRPFHFDVAEDPEVHIGOLM RSLSELQVATDSTFSLIGRGFGKVKYKGLAGSLVAARKLERRTGELOTELOLE ALISNAVRNLRLRGFCOMPTERLILVYPMANGSVASGLREROSSEPLDMPTRKR MGSARGSLYLDHDCPKIIRHDYKANILPNDSEPAVGDGILAKINDYDHTHTTA VNGTIGYIAPETLSGTGKSSKSEKTDYFGIMLELITLITGQAFDLRLANDDDMLIDRW KSLERKLEMLVDPDENNTDITVEBQLIQVALICTQSPSPRPMSPSVYRMLEGGD LAEKDEMOKEVYHODELAPHTSSTMIIDSDYNDHARELSGPR"			
BASE COUNT			
531 a 354 c 415 g 515 t			
ORIGIN			
Query Match	7.6%;	Score 84.2;	DB 5;
Best Local Similarity	52.1%;	Pred. No. 3.5e-10;	Length 1815;
Matches 188;	Conservative 0;	Mismatches 173;	Indels 0;
Gaps 0;			
Oy	399	tgcgcctgagcttgggaagcttgaacattacagatctagagctctacaacaacacat	458
Dd	138	TGATGCTTACTTGACAAATATGCGGTTCTTATGACATTGGAGCTTTCACAAATACAT	197
Oy	459	ccaaggaacatataccttcgcgaacttggaaatctgaagaatctcalcagcttggatctgta	518
Dd	198	AAGTGGACCAATTCCTGATGATCTTGGAACTGCAAAATTTGGTGAGCTTGACCTATA	257
Oy	519	caacaacaatcttaacagggaatagttcccacttccttgggaaatggaaatgcctgcgtctt	578
Dd	258	CATGATATAGCTCTCTTGACCTTATACCGACACATTAGGAAGACTTACAAAGCTTAAGATT	317
Oy	579	ttacagccttaatgacacacagcattgacgcggtccaaatccttagagcactacagcgaatcca	638
Dd	318	CTTGCGTCTCAACACACAGCCTCTCTGTCACATTCOAATGTCACGTACTATATTATAC	377
Oy	639	agccttcaaatgttctgacgtctcaagcaagatattgttggggacaatccacaacaacg	698
Dd	378	AACCTTCACACTCTGGATTATCAACACATAGCGCTATCAGGACCAAGTACGGATAATGG	437
Oy	699	acccttgcctacatctcttcaagacattgagacaacaccgagatttggaggacccga	758
Dd	438	CTCATTTTCTTTGTTTACACCTACAGTTTGTGCAATATATTGAAATTATGTGACCCGT	497
Oy	759	a 759	
Dd	498	A 498	
RESULT 11			
LOCUS	AC021198	81513 bp	DNA
DEFINITION	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete		
ACCESSION	AC021198		
VERSION	AC021198.2	GI:6957696	
KEYWORDS	HTG.		
SOURCE	Thale cress.		
ORGANISM	Arabidopsis thaliana		

REFERENCE AUTHORS	1 (bases 1 to 81513) Liu,S.X., Sakano,H., Yu,G., Lee,J.M., Lenz,C., Pham,P., Tortumt,M., Chin,C., Chlou,J., Choi,E., Chung,M., Gonzalez,A., Hongg,B., Liu,A., Vaysberg,M., Altati,H., Brooks,S., Buehler,E., Chao,Q., Conn,L., Conway,A.B., Hansen,N.F., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.
TITLE	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 81513)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE	3 (bases 1 to 81513)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	On Feb 11, 2000, this sequence version replaced g1:6693723.

FEATURES	Location/Qualifiers
source	1. 81513
	/organism="Arabidopsis thaliana"
	/cultivar="Columbia"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="F1AD7"
BASE COUNT	27779 a 13814 c 13868 g 26052 t
ORIGIN	
Query Match	7.28; Score 79.8; DB 50; Length 81513;
Best Local Similarity	53.7%; Pred. No. 3.9e-09;
Matches 165; Conservative	0; Mismatches 142; Indels 0; Gaps 0
QY	383 accctctcgacacctcgcgcctcgcgcctcggagcttgggaagcttgaacatttcaagtatctcagac 442
DB	2466 ATCTAACTGTGGGCATTCCCCGAAACTAGACACATAGAAATCCATGATCGATTAGAGT 2525
QY	443 tctcaaaaaaacatccaaagaaactataactcgaacttggaaatctgaagaactca 502
DB	2526 TGAGTAACAACAACCTACCGGTTCAATTCCTCTCTTAGGAATCTCAGAAATTGA 2585
QY	503 tcaagcttggatctgtataaacaacaacatcttaagaggtatctccaaacttcttggaaaat 562
DB	2586 CTATCTCTTATCTTTACGAAACCTATTTAACCTGCTGTCATCTCCCGGAACCTAGGCCACA 2645
QY	563 tgaagtcctgtctctttttagcgcctaatcacacacgatcgcggtcccaatcctagagc 622
DB	2646 TGGATTCATATTTGATTTACAGTAGTAATACAAACAACCTTACCGGTTCAATTCTTCTT 2705
QY	623 actcagcgaatcccaagcctttaaagcttgaagcttgaagctcgaagcaatgatttggtaga 682
DB	2706 CCTTGGAATCTCAAGAACTTGACGATCTTATCTTACCTGATTTATTTAACTGGTG 2765
QY	683 caatccc 689
DB	2766 TCATTC 2772











QY 624 ctcacggaatcccaagctttaaagtgtgacgtctcaagcaatgattgttgaggac 683  
 DB 84538 ATTAGAGAGAAACGGTAAACTGATCTCTTCAATAAGCTGACCGGAC 84479  
 QY 684 aatcccaacaaagcagccttgctcaca 712  
 DB 84478 TTTACCGCGGACATGTCTCGGATACA 84450

RESULT 14  
 INU77888 5033 bp DNA PLN 23-NOV-1996  
 LOCUS Ipomoea nil receptor-like protein kinase (lnrpk1) gene, complete  
 DEFINITION  
 CDS  
 U77888.1 GI:1684912  
 KEYWORDS  
 SOURCE Japanese morning glory.  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Asteridae; euasterids I; Solanales; Convolvulaceae;  
 Ipomoea.  
 REFERENCE  
 1 (bases 1 to 5033)  
 Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajeevan,M.S.  
 Identification and Preliminary Characterization of an Unusual  
 Leucine-rich Repeat Receptor-Like Protein Kinase from Morning Glory  
 (Ipomoea nil)  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 5033)  
 Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajeevan,M.S.  
 Direct Submission  
 TITLE Submitted (12-NOV-1996) USDA, ARS, Appalachian Fruit Research  
 JOURNAL Station, 45 Wiltshire Road, Kearneysville, WV 25430, USA  
 FEATURES  
 SOURCE  
 1..5033  
 /organism="Ipomoea nil"  
 /cultivar="Violet"  
 /db\_xref="taxon:35883"  
 /tissue="young leaf"  
 237..251  
 TATA\_signal  
 gene  
 681..695  
 1144..1456  
 /gene="lnrpk1"  
 join(1144..4108,4202..4566)  
 /gene="lnrpk1"  
 /note="leucine-rich repeat; protein kinase homolog"  
 /codon\_start=1  
 /product="receptor-like protein kinase"  
 /protein\_id="AA036558.1"  
 /db\_xref="GI:1684913"  
 /translation="MKAVNFTLLSLCTSTSTIAAPALNSDGAALLSLFRHWTPISD  
 ITQSNWASDTPSCWMLGVEDRQRFVTLNLSVGISSGEGPISHLKIKKVLISGN  
 GFPGSIPQLNCSLLEHIDLSNSFTGNDPDTGALQNRNLSFNSNIGFPESL  
 LSPLELFTVFTGNGLSIPNSGNSSELTLNDNPGSPFSGNSGNTIGLEY  
 LNDNNLVTPLNTNLENTLYLDVRNNSLVGALPDVFCQDIDITSLNNQFTGL  
 PPGNGCTSLRPGAFSCALSGPISPCFQGLTKDITLYLNGKHSRIPPELCKGSM  
 IDIQOQNLGSGETPELGHMSQQLYHLYTNNISGVPSSTIKIOSLSOILQYNNL  
 SGELPVDNTELDQVLSALVLENHTGVIPDUGANSLVELDITRNMFTHLPNLS  
 OKTKRLRLGYNVLSGVSDLGCGSTLERLLENNLRGLPDEVRKQMLPEPILCS  
 NNTGPIPPSLSGNLKNVAYISLSSNOLSGSPPELISGLVLEHNLSHNLKILPSE  
 LSNCHKLSLSDLSNHLNLSISPTLSGLSTELTKLSIGENSGGIPYSLNKKLNL  
 OLUGNLADIPVGAOLALRSNLNSKLNKGQPLDGLKMLELSDVHNHLSGTL  
 RYLSTIOSLFTINLSHNLFSGPVPSLPKFNLSGSPSFGNSDLCINCPADGLAPES  
 SLIRPCNNQSTNGGGLSTGLAMIVGALFTICPLPSPADLFLCKHCVQSIATSA  
 QGDSGLNKVLENTLENDKYVIGKAGHTYKATISPRKYAVAKKLVYTKGNSY  
 SWAKLETIGVYRHNLLKLEFWLKRKRGILITTYMENSLSHLDLHEINPPPLWS  
 TRHNIAVTAIGLAVLHFDQPAIVHDIKPMNLILDSLEPHISFGALADPQDQST  
 SLPSNTVOGITGVAPENAFITVRSRDVYSGVVLIEDITRKALDSEFNETDITV  
 GWNRVMTGTGTLQIKIYDPSLDELIDSVMEQVTEALSLALCAKEVDRKRPMTADV  
 VKQLFRMRSIRSSSVRNKSK"

BASE COUNT 1365 a 1060 c 1017 g 1591 t

ORIGIN

Query Match 6.9%; Score 76.6; DB 8; Length 5033;  
 Best Local Similarity 49.1%; Pred. No. 2.3e-08;  
 Matches 202; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 285 ccaagatgagatcaactctgttaactcctgttacctgtgctacatcactgtaca 344  
 DB 1281 CCAGAGCTGGAAATGCTTCAATCCATCTCTTGTATGCTGGAGTGAATGTGACAG 1340  
 QY 345 agacaacgcgtcaactcgtgtgattgattgaaattcaaaccttcggacatctgccc 404  
 DB 1341 GAGACAAATTGTATACTCTGAACTCTCTCTTANGAATCTAGGCGAATTGGGCC 1400  
 QY 405 tgaacttggaaagcttgaacattacagatctagagctctacaacaaacacacaa 464  
 DB 1401 CGAAATCTCGCATTTGAAAGCATTTGAAGAGGTGTTCTCAGTGCAGCAATGTTCTTGG 1460  
 QY 465 aactatccttcggaacttggaaatcgaagaatcgaatcgaatcgaatcgaatcga 524  
 DB 1461 CTCAAATCTCTCCAGCTAGGCAATTCAGCTCTTGAACACATGATCTGCTCCAA 1520  
 QY 525 caactctacaggaatgattcccaattctcttgggaaattgaagctctgtctttttag 584  
 DB 1521 CAGCTTACTGTAATATCCCTGACACCTTGGAGCTTTCAGAAATTTAAGCACTTAG 1580  
 QY 585 gctaatgacacacgattgacacggtccaatccatcagagcactacgagcaatccaa 644  
 DB 1581 CTTGTTCTTAAATCTCTGATGAGCCATTTCCGAGTCTTTCATTTCAATTCACATTT 1640  
 QY 645 taaagtgtgacgtctcaagcaatgattgtgtgagcaatcccaaca 695  
 DB 1641 AGAACTGTATTATTTCACTGCAATGCTTTAATNGTTCAATCCCTTCAA 1691

RESULT 15  
 AP000815/c  
 LOCUS Oryza sativa genomic DNA, chromosome 1, clone: P0003H10.  
 DEFINITION  
 REFERENCE AP000815  
 AUTHORS  
 TITLE Direct Submission  
 JOURNAL  
 SUBMITTED (30-NOV-1999) to the DDBJ/EMBL/GenBank databases, Takuji  
 Sasaki, National Institute of Agrobiological Resources, Rice Genome  
 Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail: tsasaki@labr.affrc.go.jp)  
 URL: http://www.dna.affrc.go.jp:82/, Tel: 81-298-38-7441,  
 Fax: 81-298-38-7468)  
 The orientation of the sequence is from T7 to SP6 of the PAC clone.  
 Genes were predicted from the integrated results of the  
 following: GENSCAN 1.0, BLASTN 2.0, BLASTX 2.0 as well as  
 SplicePredictor (October 1999 version). The genomic sequence was  
 searched against the non-redundant database NRP (PIR, SWISSPROT,  
 GENPEPT, PDB) from MARF DNA bank and the cDNA sequence database at  
 RGP. Protein similarities of the coding regions were searched  
 against NRP with BLASTP 2.0. ESTs represent the identified cDNA  
 sequences using BLASTN 2.0 with the corresponding DDBJ accession  
 no. and RGP clone ID.  
 Detailed information on assemble quality together with annotation  
 of this entry at http://www.dna.affrc.go.jp:82/genomicdata/genome1

ORIGIN



FEATURES		ished.html.
source		Location/Qualifiers
CDS	1..142418	/organism="Oryza sativa"
		/cultivar="Nipponbare"
		/db_xref="taxon:4530"
		/chromosome="1"
		/clone="P0003H10"
CDS	join(2901..2951,3129..3317,3521..3850,4405..4638,5069..5509)	/note="Similar to Arabidopsis thaliana DNA chromosome 4, BAC clone F22K18 (M035356)"
		/codon_start=1
		/protein_id="BAA87822.1"
		/db_xref="GI:6498419"
		/translation="MGEPNQRRLRYWOGSNKFLCGRLIFGPDAGSLFSTVLIVAPLVGICCCITKNSISSEKQVGLPVLIATIVGLADLFLMTSSRDGIIIPRNAPPEGSGGDEGVAGDYTPSAEWVTAAPRLPRKDYVNGCVYKVCYDCILLYRPPRASHSICNNKVRKRDHRCPPVGGCIGLRNRPFLFISTILCYVYVPSMINTVAHDKNDGSLKSMAGEPLSVLVITYTFVSVFVGLVFHLITLSTQTYENFRIYDKKENPVRGAISNTAEVFCAGIPPSNMNFWAPPLEPDPVSGQLPRNGADLTGGVKEKVDLEWRNGGILPAILRGIDYDEMKNDVSHIDRGAAPADPFVAGRWNHDECESTPTAVSHVNSERL"
CDS	join(9165..9635,9735..9799,9992..10319,10422..10520,10743..11006)	/note="ESTs D47452(S12946), C23573(S10086), C19173(E10057), A0081294(E10057), A0032453(S10086) correspond to a region of the predicted gene.; Similar to PNL134. (U37437)"
		/codon_start=1
		/protein_id="BAA87823.1"
		/db_xref="GI:6498420"
		/translation="MASSAMELSLNPAARMGLSAKPRVYSSRRIVRFAVSAAPPAARCPKPKRGTEIOETLLTPRYTDDDEMRFPNAEINKQLNOEFPDLLOEFTDVAIFVRNPERKADAKMEGLDLOIYETLERSCIAFSGLIKELGRKLTNPNVAEITSLMSRDEARHAGFLNGSLDNFALDGLCTAKRTYFKPKFIYATYLSKIGYWRVYITIFRLKANPEOYVPIKFEYEMNCODENRGGFSAALAOQFLMDKAKLSRFECISLVYVLMYNDQRTYEEGIGDLTEFMDHVIETNRTARFPLVDLDEVEFERKLDRAVEINRKIIAIGESDIPLVKNKRIPHVAAVSEIIAATIMPEI ESGVDPAEPEPOLVY"
CDS	join(13373..13457,15836..15914,17337..17477,17560..17634,17726..17860,18400..18587,18681..18749,18826..18916,19028..19162,19323..19379,19349..19441,19450..19509,19523..19585,19589..19701,19800..19941,20057..20294,21339..21586)	/note="ESTs A0081256(C53656), A0081257(C53656), A0032236(R3784), A0081351(E61905), A0031587(E61905), A0081361(R3784) correspond to a region of the predicted gene.; Similar to alien-like protein. (AC005623)"
		/codon_start=1
		/protein_id="BAA87824.1"
		/db_xref="GI:6498421"
		/translation="MEDYGFESDDEPEDVDIENQYNSKGMVETDEGALAGFDQVVRMEPERAKMGFKALQOTVLYKLGKREKMDAYEMTLIKSAVTRYSKCINDIMDVSSASQNSFSLQEPYOTIKALEAKNERLAKTKLCKIMPDEGRSKILKELHSCQREDSDDOKGTQLEIYATLEIQYTEIKNNKLKLEYIKALSTKAI PPHRMGIIREGGKMAERQWADATDFEAFKNDENGPNRIQCLTYVLANLMESEVNPFGDEAPRYKNDPEILAMTNLIAQKNDIMEKILKSRRIMDDPIRNYIEDLKNIETQVYLKIKRYTRIRIPISOKYCKMIISNMEATILVFEHITLYTVAISGCKRYSLKLDLFLSCQRTVOLKSCALNADPMLFLENNILSVQVALSIVLMSCEKLNPERKDEBOLVSLIDNRIGOHIDVVKLLEDRLETPRSQVAAE EHLRNVQOSVREIACIGCTAGTCLHVLILMMKMGFCPELINSANSNETIATSM ARNWFPCNHCRTSKFWTGREVWFRVYIAGTTRTANIVLPSSDSVAQDSLSNQDE KESPLSTYTDOALNLOREHTTILSDVINLSSESYRLAPSSHLGHPYTTIYVDN I"
CDS	join(22475..22550,23332..23705)	/note="ESTs A0030013(E50493), A0081341(E50493) correspond to a region of the predicted gene.; Similar to O.sativa gene encoding calmodulin. (Z12828)"
		/codon_start=1
		/protein_id="BAA87825.1"
		/db_xref="GI:6498422"
		/translation="MADQLTDQIAERKEAFSLPDKDGCITTKELGTVMRSIQNPFTEALQDMINVEDADNGTIDPFPELIMARKMKMDTVESEELKEAFVFKDNGFIS
CDS	AAELRHYMTNLGKLTDEVEEMIREADVDGQQLNDEYFVMAK"	
	complement(join(24778..24897,25615..25722,25808..25893,26090..26177,26435..26638,27519..27581,27708..27778,28329..28389,28512..28552,29058..29225,30476..31136))	
	/note="hypothetical protein"	
		/codon_start=1
		/protein_id="BAA87826.1"
CDS	/db_xref="GI:6498423"	/translation="MMFQSTSDARTHSGDRGSAGRIACRGTRSAAPGSGSEATTRGHALAEPPHASPRLVGRGSRSPSPSPPLPYARCDEEHRHRCSSLPBR RORRGAPPPDPASPQCDLATLGVGRMKMPAAEGVCAVEEERMAARGLRGGGGVYVAHGRKSRAREERRCGFEERSGDEGVAAREEDKAEVGRNKGRTIG TDERMRGSHKRSITLMQOTKPPARRGAVNIAHCGDDRRRCRCQAVARRQEA LPIQSLPILFIMRDLKVAKEQDQIGYAGVASCITQVAYSEYCKREHATLSL VTSRRAIADVGAIPAIGFUSQETLHMHODKEVIDALBAQDATSDIGETKSGSGM GHTSRLKNNQMSAITTVCSEFLHDTAYLEIFSLMAVSSKRYRGLSFTSODGVILA ISGCVLYQLAIPILAKVYPIPFRRVAAVLSILLSTYFPMNLVGLLEKVLINI ASLKNEARKKEMTNQISLSDAGFHGRKSTLGCSTYVKGQSSKP"
	complement(join(34468..34544,37202..37385))	
	/note="hypothetical protein"	
		/codon_start=1
		/protein_id="BAA87827.1"
CDS	/db_xref="GI:6498424"	/translation="MAAGDKAGDDAAAAAPLTVSAAGRRRCPCGCTEERCKADA GIPLVNFYIMVYCISFPPRRYRPARKICDEFVKMAHGHR"
	join(40347..40358,40468..40680)	
	/note="hypothetical protein"	
		/codon_start=1
		/protein_id="BAA87828.1"
CDS	/db_xref="GI:6498425"	/translation="MHPKHLVAAATVILVQSLPSSFSNRRRRRCRPTIATADGSTRVLPDPDGGKGFRRMRKEGG"
	complement(join(42025..42083,43481..43661))	
	/note="hypothetical protein"	
		/codon_start=1
		/protein_id="BAA87829.1"
CDS	/db_xref="GI:6498426"	/translation="MAAGDKAGDDAAAAAPLTVPAAGRRRCPCGCTERCEAHG IPVNFYIMVYCISAVNRVETEMPTSTVS"
	join(47081..47084,47194..47313,47469..47584)	
	/note="hypothetical protein"	
		/codon_start=1
		/protein_id="BAA87830.1"
CDS	/db_xref="GI:6498427"	/translation="MSRGRAPATACGSYRCYAAEVAPEBDSRYALLILRLDSP AATPPLIARRIRPHDGPFLEREDLSGPTLD"
	join(48412..48822,49072..49179,49283..49438,49876..50091,50202..50411,50488..50625,50842..50943,51105..51170,51274..51358,51454..51551)	
	/note="Similar to Arabidopsis thaliana chromosome II BAC T27A16 sequence; hypothetical protein. (AC005496)"	
		/codon_start=1
		/protein_id="BAA87831.1"
CDS	/db_xref="GI:6498428"	/translation="MVLPLSLVSCRVPSPAPRSRRFLDPGGGNGVDGAGVYFRRRLRGIVASNTSSSSSRGRHDDRHGGVGDGDGADALASVRLRLSSQADDAAEEBDEEGQPKRWIAVDFGKAFLLCNDNRMSATLIPMAEKGWNPQVGL IOSFEFADVLTQIAGIWDVTVGKLVLFQFVYVNSALALPPAKGLPLLVTR AFMGVGEKLVAPANNILSLKAVSPSESRSLAVYSGMYLGSYGLAFSLPHNMG PSVFYSGSLGVWFSTWASKAVSSPLEDGIABEKKITLSOTGGEPKELPMGLI LSKRPVVALIYSHRCNMGFTILLWPPYVNOYKRNITLESGLFCVILPMTMASTN FGGIADTVLSRGISVTVKRIKNSITFLPAFPLTOLSHIDSPPAAVILCMQSGVDA FESQGLYSNODIGPRAGVLGLSTAGVLAGVGTAAVTGILLDHGSDVDVFKVSV VLYLVGLVNNLFFSTGERIID"
	complement(join(52616..52685,53228..53299,53395..53466,53657..53800,54663..54734,56026..56234))	
	/note="ESTs C97644(C60871), A0081265(C60871) correspond to a region of the predicted gene.; Similar to L.esculentum LRP gene. (X95269)"	
		/codon_start=1
		/protein_id="BAA87832.1"
CDS	/db_xref="GI:6498429"	/translation="MAAWSPALAAVLLAAVAASNSBGDALYLRALADPRVLQ



CDS

SMDPLVNPCTWEHYVTCDRAGVTRLDGNSNLSGHLAPELGHLEHLOYLEYKNNTIO  
 GTIPAEIGSLKNIISLDLYNNNTGTIPRELGKLSIYFLINDNSLNGPIPRDLAKI  
 SSKYIVDSNNDLCSGITPISGPEHPIPLNPDKNPRLSPPELGSLATIDYNC"  
 59608..59826  
 /note="hypothetical protein"  
 /protein\_id="BAA87833.1"  
 /codon\_start=1  
 /db\_xref="GI:6498430"  
 /translation="MKLVDPHPRSHHVFYLA/SLTSLATYSLTIIVKTKGLAVPE  
 KNGTSSLTGSVSPKDRSRNRPGVNLADH"  
 complement(join(63589..63842,63844..64145))  
 /note="hypothetical protein"  
 /protein\_id="BAA87834.1"  
 /codon\_start=1  
 /db\_xref="GI:6498431"  
 /translation="MRSRAPTDLRAAGRAATLRRGRGRSRGQEEYLGTA/SLGGG  
 EGSTDEWEGRGEGGAAVPAPEGERRRRAETKLGSLIVGGRGAPHRQSGHG  
 EGGRDEEGELMLVGGALRPEEEDEEGHGRNRGRGERRRVCERERER"

CDS

Query Match 6.98; Score 75.8; DB 7; Length 142418;  
 Best Local Similarity 67.3%; Pred. No. 3.6e-08;  
 Matches 107; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Oy 433 tatctagagctctacaaacaacatccaggaactatacttcggaacttgaaatctg 492  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 53805 TACAGAGAGCTGTACAAACAAATATTCAGGGCACTATCCAGCTGAGCTCGGAGTCTG 53746  
  
 Oy 493 aagaatctcatcagcttgatctgtacacaacaatcttaagggatagttcccacttc 552  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 53745 AAGAACCTTATTACCTTGAGACTGTATACACACATTACTGGAACCATACCGAAGAA 53686  
  
 Oy 553 ttgggaaattgaagtcctgtgctcttttaaggcttaatt 591  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 53685 CTTGGCAAGCTGAGTCTCTAGTATCTCTGTGAGTAACT 53647

Search completed: June 24, 2000, 00:15:32  
 Job time: 45887 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:52:13 ; Search time 446.21 Seconds  
(without alignments)  
620.139 Million cell updates/sec

Title: US-09-180-798-22

Perfect score: 1106  
Sequence: 1 tcgacccgcgcgtccgcctcca.....aaaaaaaaaaaaaaaaaaaaa 1106

Scoring table: IDENTITY\_MUC  
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1106	100.0	1106	1	V06586	Arabidopsis thaliana
2	924	83.5	981	1	V06587	Arabidopsis thaliana
3	912.4	82.5	1063	1	V06590	Arabidopsis thaliana
4	714.2	64.6	788	1	V06588	Arabidopsis thaliana
5	330	29.8	894	1	V06589	Arabidopsis thaliana
6	191	17.3	2089	1	V06591	Arabidopsis thaliana
7	84.2	7.6	1814	1	V06571	Daucus carota SERK
8	74	6.7	6695	1	V06570	Daucus carota SERK
9	72.4	6.5	4081	1	V06585	Arabidopsis thaliana
10	70.8	6.4	3176	1	T62124	Arabidopsis thaliana
11	64	5.8	3573	1	T06307	Partial tomato pat
12	64	5.8	6471	1	T06306	Tomato pathogen re
13	61.2	5.5	3979	1	V14518	CF-5 pathogen resi
14	61.2	5.5	3979	1	V14519	CF-5 pathogen resi
15	61.2	5.5	4123	1	V14523	CF-5 pathogen resi
16	61	5.5	3293	1	X23532	Tomato Xa21 clone
17	58.6	5.3	6256	1	T31329	Rice Xa21 disease
18	58.6	5.3	7204	1	X23527	O. longistaminata X
19	57.6	5.2	3045	1	X23531	Maize Xa21 gene DT
20	57.4	5.2	3921	1	T31300	Rice Xa21 disease
21	57.4	5.2	5940	1	X23526	O. longistaminata X
22	57.4	5.2	9424	1	X23525	O. sativa Xa21 gen
23	57.4	5.2	13340	1	X23522	O. longistaminata
24	57.4	5.2	19639	1	X23524	O. longistaminata
25	56.4	5.1	1554	1	T31307	Tomato RRR gene cl
26	55.6	5.0	3541	1	V14522	CF-5 pathogen resi
27	54.8	4.9	8416	1	X23523	O. longistaminata
28	54.4	4.9	2075	1	T49435	Tomato polygalactu
29	53	4.8	1058	1	T49434	Pear polygalactu
30	51	4.6	4104	1	X07356	Arabidopsis sterol
31	50	4.5	2192	1	X23530	Maize Xa21 gene DT
32	49	4.4	5733	1	X00477	Arabidopsis thaliana
33	48.2	4.4	3842	1	X23533	Tomato Xa21 clone
34	46.2	4.2	3050	1	T06309	Tomato pathogen re

35	46.2	4.2	3089	1	T47877	Tomato pathogen re
36	41	3.7	1957	1	Q05644	Clone hns16 encodi
37	40.8	3.7	792	1	Q42596	Sequence of varian
38	40	3.6	5852	1	Q11710	Dicystosellum plas
39	39.6	3.6	19124	1	T72882	Plasmidium var-7 g
40	39.4	3.6	4590	1	N60472	Sequence encoding
41	39	3.5	2880	1	Q21450	Tomato Cf-9 cDNA.
42	39	3.5	2880	1	T06304	Tomato pathogen re
43	39	3.5	3905	1	Q91449	Tomato Cf-9 gene.
44	39	3.5	3905	1	T06303	Tomato pathogen re
45	38.6	3.5	1276	1	T42071	Human nuclear prol

## ALIGNMENTS

RESULT	1
ID	V06586
AC	V06586: standard; cDNA to mRNA; 1106 BP.
DT	03-AUG-1998 (first entry)
DE	Arabidopsis thaliana SERK LRR homologous EST clone.
KW	receptor kinase; apomictis; apomictic; seeds; production; embryos;
OS	plant breeding; leucine-rich repeat; ss.
FW	Arabidopsis thaliana.
FT	Key
FT	Location/Qualifiers
FT	CDS
FT	142..798
FT	/tag="a
FT	/note="Shows high homology to SERK"
PN	W0913427-A1.
PD	20-NOV-1997.
PF	13-MAY-1997; E02443.
PR	14-MAY-1996; GE-010044.
PA	(NOVS) NOVARTIS AG.
PI	De Vries SC, Hecht VRC, Schmidt EDL, Van Holst GJ.
DR	WPI; 98-08529/08.
DR	P-SDB; W47018.
PT	Production of apomictic seeds - useful in plant breeding
PS	Claim 28; Pages 71-73; 123pp; English.
CC	The sequence is that of an EST clone showing high homology to
CC	SERK LRR (leucine-rich repeat) sequences.
SQ	Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;
Query Match	100.0%; Score 1106; DB 1; Length 1106;
Best Local Similarity	100.0%; Pred. No. 8.9e-212;
Matches 1106; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 tcgacccgcgcgtccgcctccaacttaataaagggaagcaacgaacctaatttgcct 60
DB	1 TCGACCCGCGGTCGCTCCACTCATTAAGGGAACCAACGTAACCTTAATTGCT 60
QY	61 ttctcctcttggcagaataatttccttactctcaaatctcttcgaattccctct 120
DB	61 TTCCTCTTGTTCAGAAATTTTCCTTACTCTCAAAATTCCTTCGATTCCTCT 120
QY	121 cttaaactcgaactacatagcggtctgaactatcgctggagctcttcgcagct 180
DB	121 CTTAAACTCGAAACCTCATGCGGTCTGCAACTNCGGTGGAGCTTTCGCACT 180
QY	181 tcgttaacctaaacttaacttgatcttcacccgctgcgaagcaacccgaagagatgt 240
DB	181 TCGTTAACCTTAACCTTAGCTTGATTGATTCACCTGTCGGAAGCAAACTCGAAGGAGATCT 240
QY	241 cttaagcctcttgcggaggttggaagatcaagacatgtctctcaagctggatcca 300
DB	241 CTTCACCTCTTGCGCGGAGTTTGACAGATCCAGACCATGCTCTCAAGCTGGGATCCA 300
QY	301 actcttgtaacctgtgaactggttcacatgtcacctgttaaccaagaacgcgtcact 360
DB	301 ACGTGTGTAATCCTTGTAACCTGTAACCTGTAACCTGTAACCAAGAACGCGGTACT 360
QY	361 cgtgtgatttggaattcaaacctctctggaacatcttgccgtgagcttggaagctt 420



|||||  
Db CGTGGATTGGGAATTCAAACCTCTCTGGACATCTTGCCGCTGAGCTTGGAAAGCT 420  
QY gaacatttcagatcttaagctctcaaaaacacatccaaagacatatactctccaa 480  
Db 421 |||||  
Db 421 GAACATTTACAGTATCTAGGCTCTACAAAACACATCAAGAACTATACCTTCGGA 480  
QY 481 ctctgaatctgaagaatctcatcagcttgatctgttaacaacaatcttaacagga 540  
Db 481 CTGGAATATCGAAGATCTCATCAGCTTGATCTGTACAAACAACTCTTACAGGATA 540  
QY 541 gtcccaactctcttgagaaatgaagctctggtctttttacagcttaatacaacga 600  
Db 541 GTTCCACCTCTTGGGAAATGGAAGTCTGTGCTTTTACGGCTTATATGACAAACGA 600  
QY 601 ttgacggtccaatctctaagcactcaagcgaatcccaagccttaagttgttgaagtc 660  
Db 601 TTGACCGGTCCATCTCTAGGACACTCACGGCAATCCCAAGCCTTAAATTTGTGACGTC 660  
QY 661 tcaagcaatgatttgtgtggaacatcccaacaacgagacccttgctcaatccctta 720  
Db 661 TCAAGCATATTTGTGTGGCAATCCCAACAAAGGACCCCTTGTCTACATCTCTTAA 720  
QY 721 caagaacttgaagaacacccgagatctgagggacggaattactcggtcttcaagctac 780  
Db 721 CAGAACTTTAGAACAAACCCGAGATTGAGGGACCGGAATTCCTGCTTTCGACAGCTAC 780  
QY 781 gacactaactgcactggaacacactgcaaaaactgaaatgaaagattgagggagtgacc 840  
Db 781 GACACTTATGACACTGGAACAACTGGCAAAACCTGAAATGAAATGGGGGGTGACC 840  
QY 841 ttgtaagaacactccacattatcaaatatcacatctatgtaaatgaatatata 900  
Db 841 TTGTAAGAACAACCTCCACCTTATCAAAATATCAATCTATATGTAATGAATATATA 900  
QY 901 tgaagtaaaaaaataaatagaagaatcgaaatcggtlaatatcatctggtctcaattgag 960  
Db 901 TGTAGTAAAAACAATAAATGAAAGATCGAATCGGAATATCATCTGGCTCAATTTGAG 960  
QY 961 aacttcgaagtcgtatgtataaattcttaaatgagatttcgtctctgtatgttcggt 1020  
Db 961 AACTTCGAGGTCTGTATGTAATTAATTTCTAAATCGAATTTCTGCTTCTAAATGTTCCGT 1020  
QY 1021 tgtggaactctgaagaatcaattgtatgtatgtatgtaacagttgcttcgctgtct 1080  
Db 1021 TGTGGGATTTCTGAGAACTAATTGTGATGTGATGTGATGATCAAGTGTCTGCTTGTCT 1080  
QY 1081 gcaaaaaaataaataaataaataa 1106  
Db 1081 GCAAAAAAAAAAAAAAAAAAAAAA 1106

RESULT 2  
ID V06587 standard: cDNA to mRNA: 981 BP.  
AC V06587;  
DE 03-AUG-1998 (first entry)  
DR Arabidopsis thaliana SERK LRR homologous EST clone.  
KM receptor kinase; apomictic; apomictic; seeds; production; embryos;  
KW plant breeding; leucine-rich repeat; ss.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 104..760  
FT /tag- a /note- "shows high homology to SERK"  
FT /tag- a /note- "shows high homology to SERK"  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
P-PSDB: W47019.

PT Production of apomictic seeds - useful in plant breeding  
PS Claim 28; Pages 75-77; 123pp; English.  
CC The sequence is that of an EST clone showing high homology to  
CC SRK LRR (leucine-rich repeat) sequences.  
SQ Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;  
Query Match 83.5%; Score 924; DB 1; Length 981;  
Best Local Similarity 98.8%; Pred. No. 1.3e-225;  
Matches 952; Conservative 0; Mismatches 10; Indels 2; Gaps 2;  
QY 51 taatttgccttctctctcttctgttgcgaataatttcccttctctcaattcccttttc 110  
Db 13 TTAGTTGCTTTTCTCTCTTTGTTGAAAAATTTCCCTTATCTCAATTCCTTTTCG 72  
QY 111 attccctccttaaacctccgaagctcacatgagcttcgaactatcagttgagagct 170  
Db 73 ATTTCCTCTTAAACCTCCGAAGCTCAGATGGGCTTCGAAACTATCGGTGGAGCT 132  
QY 171 ctctgagcttctgtaaccttaaccttaaccttgatctcaactggtcgaagaactcga 230  
Db 133 CTTCGACGCTTGGTAACTTAACCTTAACCTTAACCTTGAATTCACCTGGTCAAGCAACTCCGA 192  
QY 231 agagagatgctctctacgctctcgcgaggtttgacagatccagaccatgctccagag 290  
Db 193 AGGAGATGCTCTTACGCTCTTGGCCGAGTTGACAGATCCAGACATGCTCTCCAGAG 252  
QY 291 ctggatcccaactcttctgttaacctgttaccgtggtcagatccactgtgaacagaa 350  
Db 253 CTGGGATCCAACTCTGTATATCTTGAATCTTGTACCTGGTCCATGCTCACTGTAACCAAGACA 312  
QY 351 ccggtgactcgtgtgatttgggaattcaaacctctcttgacatcttgcctgagagct 410  
Db 313 CCGGTACTCGTGTGATTTGGGAATTCAAACCTCTCTGGACATCTGCTCTGAGCT 372  
QY 411 tgggaagcttgaacatttacaagatctagaagctctcaaaaacaaacatccaaagaaat 470  
Db 373 TGGGAACCTTGAACATTTACAGTATCTAGAGCTCTCAAAAACAACTCAAGAACTAT 432  
QY 471 aacttcgaacttggaaatctgaagaatctcatcagcttggatctgtataaacaacatct 530  
Db 433 ACCTTCGGAACCTTGAATATGGAAGATCTCACTACCTTGGATCTGTAAACAACTAT 492  
QY 531 tacagagatagttcccaactctcttggaataattgaagctctggtcttcttaacgttaa 590  
Db 493 TACAGGATAGTTCACCTCTTTGGGAAATGGAAGTCTGCTGTTTAAOGGTTAA 552  
QY 591 tgaacacgatttgcaggtcccaat-ectagagcactcaacggaatcccaagccttaag 649  
Db 553 TGACAAACGATGACCGGTCCTTAAGCAGCTCACGGCAATCCCAAGCC-TTAAAG 611  
QY 650 ttgttgcgctctcaagaatgatttgtgtggaacaatcccaacaacggaaccttgcctc 709  
Db 612 TTGTTGACGTCTCAAGCAATGATTTGTGTGGAACAATCCCAACAACGGAACCTTTGCTC 671  
QY 710 acattccttcaagaacttgaagaacaacccgaagattggaaggacccgaattactcgtc 769  
Db 672 ACATTCTCTTACAGAACTTGAAGAACCCGAGATTGGAGGACCGGAATTACTGGTTC 731  
QY 770 ttgcaagctcgaagaacttgaacactcgaagaacactcgaagaacccgaagaatgaagaat 829  
Db 732 TTGCAAGCTACGACACTTACCTGCACTGGAACAACTGGCAAAACCTGGAATAAGAAATTT 791  
QY 830 gggggtgaccttgaagaacacttcaacacttatacaatatacatatcatatgta 889  
Db 792 GGGGGGTGACCTGTGAAGAACACTTACCACTTAAATATCAATATCATATATATAT 851  
QY 890 aagtatatatgtagtaaaaaaataaataaataaataaataaataaataaataaataaata 949  
Db 852 AAGTATATATATGTATGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 911  
QY 950 tctcaatgagaactcgaagctgtgtatgtaaatcttaaatgagcgaatttcgcttaccg 1009  
|||||



Db 912 TCCTCAATGGAACCTCGAGCTCTGTATGTAATAATTCTTAATGCGATTTCGCCCTAAT 971  
QY 1010 taat 1013  
Db 972 TACT 975

## RESULT 3

V06590  
ID V06590 standard: cDNA to mRNA; 1063 BP.  
AC V06590;  
DT 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK LRR homologous EST clone.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KW plant breeding; leucine-rich repeat; ss.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 106..162  
FT /note= "shows high homology to SERK"  
FN W09743427-A1.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
P-PSDB: W47022.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 28; Pages 86-88; 123pp; English.  
CC The sequence is that of an EST clone showing high homology to  
CC SERK LRR (leucine-rich repeat) sequences.  
SQ Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;

Query Match 82.5%; Score 912.4; DB 1; Length 1063;  
Best Local Similarity 95.0%; Pred. No. 1.2e-222;  
Matches 1011; Conservative 0; Mismatches 36; Indels 17; Gaps 6;

QY 39 ccaagtaacccattatgtca 98  
Db 14 CCGAGGAACCCCTAATTTTGGTCTCTCATC-TTCTTAGAAAAAT-----ACTCA 62  
QY 99 aattccttcgattccctctctctctctctctctctctctctctctctctctctca 158  
Db 63 AATTCCTATTAGATTACT 122  
QY 159 tcgctggagagctcttcgagctctctctctctctctctctctctctctctctctctctca 218  
Db 123 TCGGTGGAGGCTCTCGAGCT 182  
QY 219 agcaactcgaagagagatgtctctctctctctctctctctctctctctctctctctca 278  
Db 183 AGCAAACTCCGAAGAGATGCT 242  
QY 279 tgcctcgaagagctggagctctctctctctctctctctctctctctctctctctctca 338  
Db 243 TGTCTCCAGAGCTGGGATCCAACTCTGTAACTCTGTAACTCTGTAACTCTGTAACTCTGT 302  
QY 339 taaccaagacaacgcgctcactcgttgatattggaattccaacctctctgaactc 398  
Db 303 TAAACCAAGACAACGCGCTCACTCGATGGGATTTGGGAATTCAAACCTCTCTGACATCT 362  
QY 399 tgcgctgaagctggagagcttgaacattcagatctcagagctctcaaaaaaacat 458  
Db 363 TCGCGCTGAGCTGGGAGCTTGAACATTACATATCTAGAGCTCTACAAAAACAACAT 422  
QY 459 ccaagaaactataccttcgaacttgaacatcgaagatctcagcttgatctgta 518  
Db 423 CCAAGGAAGTAACCTTCGAACTTGGAAATCTGAAGATCTCATCAAGCTTGGATCTGTA 482  
QY 519 caaacaacattctacagagatagttcccaacttctcttggaataattgaagctctgctct 578

## RESULT 4

V06588  
ID V06588 standard: cDNA to mRNA; 788 BP.  
AC V06588;  
DT 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK LRR homologous EST clone.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KW plant breeding; leucine-rich repeat; ss.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 2..664  
FT /note= "shows high homology to SERK"  
FN W09743427-A1.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
P-PSDB: W47020.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 28; Pages 79-80; 123pp; English.  
CC The sequence is that of an EST clone showing high homology to  
CC SERK LRR (leucine-rich repeat) sequences.  
SQ Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;

Query Match 64.6%; Score 714.2; DB 1; Length 788;  
Best Local Similarity 96.2%; Pred. No. 2.2e-172;  
Matches 753; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

QY 139 cacatgagcttcgaacatcgttgagagctcttcgagcttcgtaaccttaacctta 198







KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 KW plant breeding; ds.  
 OS Arabidopsis thaliana.  
 FH Key Location/Qualifiers  
 FT CDS 195..2072  
 FT /tag= a  
 FT /product= SERK protein  
 PN M09743427-A1.  
 PD 20-NOV-1997.  
 PE 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 DR WPI: 98-086529/08.  
 DR P-PSDB; W47023.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 27; Pages 91-95; 123pp; English.  
 CC The sequence is that encoding SERK, a putative receptor kinase.  
 CC It may be used as part of a method of producing apomictic seeds  
 CC comprising: (a) transforming plant material with a nucleotide  
 CC sequence encoding a protein which in active form in a cell or  
 CC cell membrane renders the cell embryogenic; (b) regenerating  
 CC the transformed material into plants or carpel-containing  
 CC plant parts; and (c) expressing the sequence in the vicinity  
 CC of the embryo sac. The apomictic seeds and embryos thus produced  
 CC can be developed into plant progeny. This is useful in plant  
 CC breeding programs. Controllable and reproducible apomixis provides  
 CC many advantages in plant improvement and cultivar development in  
 CC the case that sexual plants are available as crosses with the  
 CC apomictic plant. Apomixis provides for true-breeding, seed  
 CC propagated hybrids and could shorten and simplify the breeding  
 CC process so that selfing and progeny testing to produce and/or  
 CC stabilise a desirable gene combination could be eliminated.  
 CC Apomixis allows plant breeders to develop cultivars with  
 CC specific stable traits for such characteristics as height,  
 CC seed and forage quality and maturity.  
 CC Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;  
 SQ

Query Match 17.3%; Score 191; DB 1; Length 2089;  
 Best Local Similarity 60.5%; Pred. No. 1,9e-39;  
 Matches 314; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 220 gcaaacccgaagagatgctctcttaagctcttcgcgcgaggttgacagatccagacat 279  
 DB 270 GCTATTGTGAAGAGTGAGCTTTCATCTTGAAGGCTTACTGATCCAAACAT 329  
 QY 280 gtcccccagagctggtatccaactctgttaactctgttaccgttccatgttacctgt 339  
 DB 330 GTCTTGCAAGCTGGAGTCTACGCTAGTGAATCTTGCAATGTTCCATGTCATCTGC 389  
 QY 340 aaccagaacaaccgcgtacactcgtgtgatttggaatcaaacctctgacatctt 399  
 DB 390 AACACAGAGACAGTGTATAGATGATTGGGAAATGAGATGATATCGGCCATTTA 449  
 QY 400 ggcgtctgagctgggaagcttgaacattacagtatctagagctctcaaaaacatc 459  
 DB 450 GTTCCAGAGCTTGCTGTCTCAAGAAATTTGCAATTTGGAGCTTTACAGTAACATA 509  
 QY 460 caaggaactatacttcggaacttgaatctgaagaatccatccatccatcttgatcttac 519  
 DB 510 ACTGGCCGATTCCTAGTAATCTTGAATCTTGACAACTTAGAGATTGGATTTGAT 569  
 QY 520 aacaacaatcttaaggagtagttccacattcttgggaaaattgaaagctctgctt 579  
 DB 570 TTAACAGCTTCTCCGGTCTATTCGGAATCATGGGAAAGCTTTGAAAGCTAGATT 629  
 QY 580 ttaagccttaatgaacaacgattgacgcgttgcacatccatccagcactcaagcacc 639  
 DB 630 CTCGCGCTTAACAACAACAGCTCTACGTGGTCAATTCCTATGTCACGACCAATATTA 689  
 QY 640 gcccttaagtggtgagctctcaagcaatgattgtgtggacaatcccaacaacg 699

DB 690 ACCCTTAACCTGTAGTCTATCAATAAACAAGACTCTCTGCTCACTTCTGACATGCG 749  
 QY 700 cccctgtcacatctctcttacagaaattgagaaacac 738  
 DB 750 TCCCTTCACCTCTCAACACCACATCACTTGGCAATTAAC 788

RESULT 7  
 ID V06571 standard; cDNA; 1814 BP.  
 AC V06571;  
 DE 03-AUG-1998 (first entry)  
 DE Daucus carota SERK gene  
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 OS Daucus carota.  
 FH Key Location/Qualifiers  
 FT CDS 94..1755  
 FT /tag= a  
 FT /product= SERK protein  
 PN M09743427-A1.  
 PD 20-NOV-1997.  
 PE 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 DR WPI: 98-086529/08.  
 DR P-PSDB; W47013.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 21; Pages 47-51; 123pp; English.  
 CC The sequence is that encoding SERK, a putative receptor kinase.  
 CC It may be used as part of a method of producing apomictic seeds  
 CC comprising: (a) transforming plant material with a nucleotide  
 CC sequence encoding a protein which in active form in a cell or  
 CC cell membrane renders the cell embryogenic; (b) regenerating  
 CC the transformed material into plants or carpel-containing  
 CC plant parts; and (c) expressing the sequence in the vicinity  
 CC of the embryo sac. The apomictic seeds and embryos thus produced  
 CC can be developed into plant progeny. This is useful in plant  
 CC breeding programs. Controllable and reproducible apomixis provides  
 CC many advantages in plant improvement and cultivar development in  
 CC the case that sexual plants are available as crosses with the  
 CC apomictic plant. Apomixis provides for true-breeding, seed  
 CC propagated hybrids and could shorten and simplify the breeding  
 CC process so that selfing and progeny testing to produce and/or  
 CC stabilise a desirable gene combination could be eliminated.  
 CC Apomixis allows plant breeders to develop cultivars with  
 CC specific stable traits for such characteristics as height,  
 CC seed and forage quality and maturity.  
 CC Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;  
 SQ

Query Match 7.6%; Score 84.2; DB 1; Length 1814;  
 Best Local Similarity 52.1%; Pred. No. 2,3e-12;  
 Matches 188; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 339 tgcgcctgagcttgggaagcttgaacattacaglatctagagcttcaaaaacacat 458  
 DB 138 TGATGCTTACCTTGACAAATATGGGCTTATGACATTTGAGCTTTACAGCATATACAT 197  
 QY 459 ccaaggaactatacttcggaacttgaatctgaagaatccatccatccatcttgatctt 518  
 DB 138 AAGTGACCAATTCCTAGTATCTTGGAAATCTGACAAATTTGGTGAAGCTTGAACCTTA 257  
 QY 519 caacaacaatcttaaggagtagttccacattcttgggaaaattgaaagctctgctt 578  
 DB 258 CATGAATGCTTCTCTGACCTATACCGACACTTGAAGAACTTAACAAGGCTAGATT 317  
 QY 579 ttaagccttaatgaacaacgattgacgcgttgcacatccatccagcactcaagcacc 638  
 DB 318 CTTGCGCTTAACAACAACAGCTCTCTGCGCAATTCGAATTCATGCTACGACCAATATTA 377  
 QY 639 agccttaagtggtgagctctcaagcaatgattgtgtggacaatcccaacaacg 698







P1 De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ,  
PI WPI: 98-086529/08.  
DR P-PSDB: W47017.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 26; Pages 64-67; 125pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilize a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
CC Sequence 4081 bp; 1120 A; 770 C; 785 G; 1406 T;  
CQ

Query Match	6.58;	Score 72.4;	DB 1;	Length 4081;
Best Local Similarity	70.38;	Pred. No. 3e-09;		
Matches 97; Conservative	0;	Mismatches 41;	Indels 0;	Gaps 0;

Accession	Sequence	Position
OY 231	agagaaagcctcctcaagctcttcgcggagtttgaaagctcagaccatgctctctccaa	290
OY 1794	AGGGAAGCGCTTGACATCTTTGAGGGTTACTCTAAGTTGATTCACCAAAAGTCTTTCAGG	185
OY 291	cttgagatcccaactcttttaactctcttacttggtccatgtaactgtgaacacagaca	350
Db 1854	CTGGAAATCCACGCTAAAGCAATCTTCCACATGTTCCATGATCTTCACACACAGAA	191
OY 351	ccgcgtactactcggtgtga	368
Db 1914	CAGTGTCTAATAGAGAGTGA	1931

RESULT	10
ID	T62124
AC	T62124 standard; cDNA to mRNA; 3176 BP.
DT	T62124;
DE	10-JUN-1997 (first entry)
DE	Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
KW	Plant; morphogenesis; regulation; short; stem; alteration;
KW	inflorescence; extraneous; gene; expression; transformation;
OS	increase; control; form; length; ds.
OS	Arabidopsis thaliana.
FT	Key
FT	Location/Qualifiers
FT	51..2981
FT	cds
FT	/*tag= a
FT	/note="plant morphogenesis regulatory protein"
PN	
PN	U09056382-A.
PD	04-MAR-1997.
PF	24-AUG-1995; 216187.
PF	24-AUG-1995; JP-216187.
PA	(MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA	(CHIK-) ZH CHIKYU KANKYO SANGYO GIJUTSU KENKYU.
DR	WPI; 97-206629/19.
DR	P-PSDB; W13308.
PT	DNA encoding plant morphogenesis regulatory protein - useful to
PT	yield plants with short stems or altered inflorescence
PS	Claim 1; Pages 6-10; 17pp; Japanese.
CC	The present sequence encodes an Arabidopsis thaliana plant
CC	morphogenesis regulatory protein (MRP), which can be used to yield
CC	a plant with, e.g. short stems or altered inflorescence. The MRP
CC	acts on a plant at a specific site for a specific period, and can

CC therefore be used to regulate extraneous gene expression in a  
CC plant. The MRP's cDNA or genomic DNA can be used to transform a  
CC plant to increase its MRP expression, and therefore control the  
CC form (particularly stem length) of the plant.

50 Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match	6.4%;	Score 70.8;	DB 1;	Length 3176;
Best Local Similarity	52.0%;	Pred. No. 7.1e-09;		
Matches 159;	Conservative	0;	Mismatches 147;	Indels 0;
			Gaps	0;

OY 390 tggacatctggscctbaagcttgggaagcttaacattacagatctcagagcctaca 449  
 Db 1226 TGGCACATACCCCGAGCATTTCAAAGCTAGAAATATGAGCTTACCTTAATCTCTCCAG 1285  
 OY 450 aaacacatccagaagaaactaactcttcggaacttggaaatctgaagaatctcatcagctt 509  
 Db 1286 CAACAAATCAAAAGTCCATCCCGGTAGCTATCTCTACGGTACTTAGATACATT 1345  
 OY 510 gacatctacaacaacaacattacaggaagttcccaattctcttgggaatatgaatc 569  
 Db 1346 GGATCTTTCCACACACAAAGTAAATATGAGATCATCTCTTCCCTTGGTGAATTTGGACCA 1405  
 OY 570 tctggtcttttaagagcttaatgacaacagatgacaggtccaatcctagagcactcag 629  
 Db 1406 TCTTCTCAAGATGAATCTGTGTAATATATTAATGGTGTAGTTCCAGGGGACTTTGG 1465  
 OY 630 gcaatcccaagccttaagaattgttagcgtccaagcaatgaattgttggtagacatccc 689  
 Db 1466 AAATCTAAGAAAGCATCATGGAATATGATCTTTCAATATATGATATCTCTGGGCCAATTC 1525  
 OY 690 aacaaa 695  
 Db 1526 AGAAGA 1531

RESULT 11

706307

ID 706307 standard; cDNA; 3573 BP.

AC 706307;

DT 14-APR-1996 (first entry)

DE Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.

KW Partial tomato pathogen resistance; Cf-2.2; tomato; C.fulvum; Avr 4; Avr 9; fungal;

KW leaf mould; variegation; ss.

OS Lycopersicon esculentum.

OS W0553564-A2.

PN 23-NOV-1995.

PD 11-MAY-1995; G01075.

PF 11-MAY-1994; GB-009394.

PR 23-DEC-1994; WO-G02812.

PR 31-MAR-1995; GB-006658.

PR 07-APR-1995; GB-007232.

PA (GATS-) GATSBY CHARITABLE FOUND.

PI Hammond-Kosack KE, Jones DA, Jones JDS;

DR WPI; 96-010949/01.

DR P-PSDB; R85299.

PT Increasing plant pathogen resistance by induction of variegation -

PT may lead to acquired resistance to a broad range of pathogens.

PS Claim 9; Page 85-87; 13pp; English.

CC 706307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone

CC In a new method this gene is expressed highly in genetic constructs

CC which may be used to impart a broad range of pathogen resistance, by

CC induction of variegation, to transgenic plants (or parts or propagules

CC of plants) containing such constructs. Cf-2.2 imparts resistance to the

CC disease caused by the leaf mould fungal pathogen Cladosporium fulvum.

CC C.fulvum contains avirulence (Avr) genes that confer recognition by

CC plants containing Cf genes, leading to the activation of host defence

CC mechanisms to attack the disease.

SO Sequence 3573 BP; 1032 A; 654 C; 664 G; 1223 T;

Query Match	Score	DB 1;	Length
Best Local Similarity	52.28;	Pred. NO.	3.9e-07;
	5.88;		



Matches 142; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

```
OY 349 aaccgcgtacgtgtgtggaatttgggaattcaacctcttggacatttgcgcctgag 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 AAAAAGCTGTGAGTGTGAATCTGTAAATATACACGCTTCTGCTCTATCTGCTTCA 1054
OY 409 ctgggaagcttgaacattacagatctagagctctacaacacacacacacaggaact 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 TTGGGGAAATCGAACAACCTGTCTATGTTGATCTTACATATACACGCTTCTGCTCT 1114
OY 469 atacccttcgaacttgaacatctgaagaatctcatcagcttggatctgtacaacaacat 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 ATTCTCTCTTCATTGGGGAATCTGAACAACCTGTCTATGTTGATCTTACATATATCAG 1174
OY 529 cttaacaggagtagtcccaacttcttgggaataattgaagctctgtcttcttaagcgtt 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1175 CTTTCTGGCTCTATCTGCTCTCATTTGGGGAATCTGAACAACCTGTGAGTTGATCTTC 1234
OY 589 aatgacaacgcatgacgggtccaatcctaga 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1235 TACAAATATCAGCTTTCTGCTCTATTCCTGA 1266
```

## RESULT 12

T06306  
ID T06306 standard; DNA; 6471 BP.

```
AC T06306;
DT 14-APR-1996 (first entry)
DE Tomato pathogen resistance gene Cf-2.1.
KW Pathogen resistant; Cf-2.1; tomato; C.fulvum; AVR 4; AVR 9; fungal;
KM leaf mould; variegation; ds.
OS Lycopersicon esculentum.
FH Key Location/Qualifiers
FT 5'utr 1..1676
    /*tag- a
    1677..5014
    /*tag- b
    /product- immature_Cf-2.1_protein
FT signal_peptide 1677..1745
    /*tag- c
    1746..5011
    /*tag- d
    5015..6471
    /*tag- e
FT 3'utr
    /*tag- e
PN MO9531564-A2.
PD 23-NOV-1995.
PF 11-MAY-1995; G01075.
PR 11-MAY-1994; GB-009394.
PR 23-DEC-1994; MO-G02812.
PR 31-MAR-1995; GB-006658.
PR 07-APR-1995; GB-007232.
PA (GATS-) GATSBY CHARITABLE FOUND.
PI Hammond-Kosack KE, Jones DA, Jones JDG;
DR WPI: 96-010949/01.
DR P-PSDB: R85298.
PT Increasing plant pathogen resistance by induction of variegation -
    PT may lead to acquired resistance to a broad range of pathogens.
PS Claim 9; Page 80-83; 131pp; English.
CC T06306 is the tomato pathogen resistance gene Cf-2.1. In a new method
CC this gene is expressed highly in genetic constructs which may be used
CC to impart a broad range of pathogen resistance, by induction of
CC variegation, to transgenic plants (or parts or propagules of plants)
CC containing such constructs. Cf-2.1 imparts resistance to the disease
CC caused by the leaf mould fungus Cladosporium fulvum.
CC C.fulvum contains avirulence (Avr) genes that confer recognition by
CC plants containing Cf-genes, leading to the activation of host
CC defence mechanisms to attack the disease.
SQ Sequence 6471 BP; 2073 A; 1106 C; 1122 G; 2170 T;
```

Query Match 5.8%; Score 64; DB 1; Length 6471;  
Best Local Similarity 52.2%; Pred. No. 4.8e-07;  
Matches 142; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

OY 349 aaccgcgtacgtgtgtggaatttgggaattcaacctcttggacatttgcgcctgag 408

Db 2679 AAAAAGCTGTGAGTGTGAATCTGTAAATATACACGCTTCTGCTCTATCTGCTTCA 2738

OY 409 ctgggaagcttgaacattacagatctagagctctacaacacacacacacaggaact 468

Db 2739 TTGGGGAAATCGAACAACCTGTCTATGTTGATCTTACATATACACGCTTCTGCTCT 2798

OY 469 atacccttcgaacttgaacatctgaagaatctcatcagcttggatctgtacaacaacat 528

Db 2799 ATTCTCTCTTCATTGGGGAATCTGAACAACCTGTCTATGTTGATCTTACATATATCAG 2858

OY 529 cttaacaggagtagtcccaacttcttgggaataattgaagctctgtcttcttaagcgtt 588

Db 2859 CTTTCTGGCTCTATCTGCTCTCATTTGGGGAATCTGAACAACCTGTGAGTTGATCTTC 2918

OY 589 aatgacaacgcatgacgggtccaatcctaga 620

Db 2919 TACAAATATCAGCTTTCTGCTCTATTCCTGA 2950

## RESULT 13

V14518  
ID V14518 standard; DNA; 3979 BP.

```
AC V14518;
DT 20-MAY-1998 (first entry)
DE Cf-5 pathogen resistance gene variant #1.
KW Tomato; Cf-5 pathogen resistance gene; Cladosporium fulvum-5;
KM tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon plimifolium.
FH Key Location/Qualifiers
FT CDS 653..3560
    /*tag- a
    /product- Cf-5 pathogen resistance gene
FT FT
PN MO9743429-A1.
PD 20-NOV-1997.
PF 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1995; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR WPI: 98-008895/01.
DR P-PSDB: W41309.
PT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT for production of transgenic plants resistant to pathogens e.g.
PT tomato leaf mould C. fulvum in tomatoes
PS Claim 3; Fig 1a; 7pp; English.
CC This sequence is an example of the polynucleotide of the invention, and
CC is able to confer pathogen resistance on a plant. It is one of two tomato
CC Cf-5 gene variants, which offer resistance against the pathogen
CC Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC the gene into plant cells and regenerating plants from the cells;
CC asexually or sexually produced offspring can also be subsequently
CC produced. Expression of the gene in plant cells can confer pathogen
CC resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC Oligonucleotides with sequences complementary to the gene or fragments
CC of it, are useful in anti-sense techniques to reduce gene expression. The
CC nucleic acids/polynucleotides are useful as hybridisation probes to
CC identify other genes/fragments conferring pathogen resistance on plants
CC e.g. Phytophthora resistance in potatoes. Homologies between Cf-5 and
CC Cf-9 may be used to identify further resistance genes of this class.
SQ Sequence 3979 BP; 1217 A; 717 C; 698 G; 1347 T;
```

Query Match 5.5%; Score 61.2; DB 1; Length 3979;  
Best Local Similarity 48.8%; Pred. No. 2.1e-06;  
Matches 165; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

OY 356 taactcgtgtgatttgggaattcaaacctcttggacatttgcgcctgagcttggga 415

Db 2023 TTACGAACTATTTTGGGTATATACCTCTTAATGGCTATTCCTGCTCAATGGGA 2082



Qy	416	agcttgaaacttcaagta	ctctgagctctcaaaaacaacatccaa	gnaactt	actt	475
Db	2083	ATCAAAACAACCTGTC	AGAGTTGATCTTACATATATAGCTTCT	GGCTTATTC	CG	2142
Qy	476	ccgactctggaatctga	aagatactcaatcgctctgatactg	tacaacaacaatct	taag	535
Db	2143	CTTCATTGGCAATAT	GAAATGGAATCTGGAAACCTGTTCT	CGATGATMAAGATCT	CATG	2202
Qy	536	ggatagttccacattct	tggaaaaatlaagctctgctctttta	acggctaa	tga	595
Db	2203	GGGAATTCCTCTATT	GTGTGTGCATTTTGACATCATCGAAGTGT	GTATATGTG	GACAA	2262
Qy	596	acgcattgaccggtcc	aatctctagagcaactcaacgcaatcc	aagcctttaag	tgt	655
Db	2263	ACAATTTGAAGGAAA	AGTTCGCCAATGTTGGGTAAATACATG	ATGACCTTCA	CAATTTGT	2322
Qy	656	acgtctcaagaatga	tattgtgtgggacaatcccaaca			693
Db	2323	CGATGTCATCTAAAT	ATGTTTCAGAGGAGACGCTCCCTTCA			2360

Sequence	3979 BP;	1218 A;	716 C;	698 G;	1347 T;
CC	CF-5	Phytophthora	resistance	in	potatoes. Homologies between Cf-5 and
CC	Cf-9	may be used to identify further resistance genes of this class.			
CC	Identify other genes/traits	confering	pathogen	resistance	on plants
CC	nucleic acids/poly/nucleotides	are useful	as	hybridisation	probes to
CC	of it, are useful	in anti-sense	techniques	to reduce	gene expression. The
CC	Oligonucleotides with sequences	complementary	to the	gene or	fragments
CC	resistance on a plant e.g. to	tomato	leaf	mould (C. fulvum)	in tomatoes.
CC	produced. Expression of the	gene	in	plant	cells can
CC	asexually or sexually produced	offspring	can	also	be
CC	the	gene	into	plant	cells
CC	Cladosporium fulvum-5.	Transgenic	plants	can	be
CC	CF-5	gene	variants,	which	offer
CC	is	able	to	confer	pathogen
CC	resistance	on	a	plant.	It
CC	is	one	of	two	tomato
CC	This	sequence	is	an	example
CC	of	the	polynucleotide	of	the
CC	invention,	and	is	useful	for
CC	production	of	transgenic	plants	resistant
CC	to	pathogens	e.g.		
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in
CC	tomatoes				
CC	Claim 6: Fig 1b: 75pp: English				
CC	DR	Tomato	gene	Cf-5,	confers
CC	resistance	to	Cladosporium	fulvum	-
CC	useful	for	production	of	transgenic
CC	plants	resistant	to	pathogens	e.g.
CC	tomato	leaf	mould	C. fulvum	in

Query Match	5.58;	Score 61.2;	DB 1;	Length 3979;
Best Local Similarity	48.88;	Pred. No. 2.1e-06;		
Matches 165; Conservative	0;	Mismatches 173;	Indels 0;	Gaps 0;

QY 356 tcaactcgtgtagattttggaattccaactctcttgcaactcttgccctgagcttggga 415  
 1 111 1 11111 111 11 111 1 111 1 111  
 Db 2023 TTACGACCTATTTTGGGTATATACCTCTTATGCTCTATTCCGCTCTATTGGGA 208  
 1 111 1 11111 111 11 111 1 111 1 111  
 QY 416 agcttgaacattcaagctatctagagctctcaaaaacaaacatccaagaactatacc 475

Db	2083	ATCTAAACA	CTGTCTAG	GTCTATCTT	TACAAAT	ATACAG	CTTCTG	GGCTAT	CTCTG	2142	
Qy	476	ccgactctg	aaatc	tgaaatc	tcatatc	taagctt	gtagatc	tgtaac	aacaaatc	cttaag	535
Db	2143	CTTATTTGG	AATATG	AAATCTG	CAAC	TCGTGTTT	TCAGTAT	ATACGAT	CTCATTG	2202	
Qy	536	ggatagctc	caactt	ctcttg	ggaat	tgaagc	tcttgc	cttctt	taagc	cttaagca	595
Db	2203	GGGAAAT	TCCTT	CTATTTG	TCTGCA	TAATTC	ACATTC	ATGCA	TGCAATG	TAATTC	2262
Qy	596	accgatgac	cggtgc	aatcct	ctagac	atc	atcag	atc	atcag	atc	655
Db	2263	ACAAATTTG	AAAGG	AAAGTTCC	GCAT	GTGTTGG	GTATAT	CAGTAC	CTTCAC	ATTGTT	2322
Qy	656	acgcctc	aagcaat	gatctt	gtgtg	ggacaat	atcc	caaca	693		
Db	2323	CGATGTCAT	CTAAAT	AGTTT	TAGAGG	AGAGCT	CTCCTT	CA	2360		

	RESULT	15
VI4523		
ID	VI4523 standard; DNA; 4123 BP.	
AC	VI4523;	
DT	20-May-1998 (first entry)	
DE	CF-5 pathogen resistance gene clone Hcr2-5D gene.	
KW	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;	
KM	tomato leaf mould; Phytophthora resistance; ss.	
OS	Lycopersicon esculentum L.	
PH	key	
FT	Location/Qualifiers	
FT	CDS	
FT	/tag= a	
FT	/product= CF-5 pathogen resistance gene clone Hcr2-5B	
PN	MO9743429-A1.	
PD	20-Nov-1997.	
PF	08-MAY-1997; G01249.	
PR	24-SEP-1996; GE-019624.	
PA	09-MAY-1996; GE-009681.	
PI	(INNE-) INNES CENT INNOVATIONS LTD JOHN.	
PJ	Dixon MS, Hatzilachanis K, Jones DA, Jones JDG;	
DR	Wpl; 98-008895/O1.	
DR	P-PSDB: WA1312.	
PT	Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful	
PT	for production of transgenic plants resistant to pathogens e.g.	
PS	tomato leaf mould C. fulvum in tomatoes	
PS	Disclosure; Fig 7; 75pp; English.	
CC	This sequence is a clone of the polynucleotide of the invention, which	
CC	is able to confer pathogen resistance on a plant. The polynucleotide of	
CC	the invention is one of two tomato Cf-5 gene variants, which offer	
CC	resistance against the pathogen Cladosporium fulvum-5. Transgenic plants	
CC	can be produced by incorporating the gene into plant cells and	
CC	regenerating plants from the cells; asexually or sexually produced	
CC	offspring can also be subsequently produced. Expression of the gene in	
CC	plant cells can confer pathogen resistance on a plant e.g. to tomato lea-	
CC	mould (C. fulvum) in tomatoes. Oligonucleotides with sequences	
CC	complementary to the gene or fragments of it, are useful in anti-sense	
CC	techniques to reduce gene expression. The nucleic acids/polynucleotides	
CC	are useful as hybridisation probes to identify other genes/fragments	
CC	confering pathogen resistance on plants e.g. Phytophthora resistance in	
CC	potatoes. Homologies between Cf-5 and Cf-9 may be used to identify	
CC	further resistance genes of this class.	
SQ	Sequence 4123 BP; 1253 A; 742 C; 725 G; 1403 T;	

Query Match	5.5%;	Score 61.2;	DB 1;	Length 4123;
Best Local Similarity	48.8%;	Pred. No. 2.1e-06;		
Matches 165; Conservative	0;	Mismatches 173;	Indels 0;	Gaps 0;

QY 356 tcaatcgctgtgagatttgggaataatcaaactctcbygnacalcttgccgtgatgta gga 415  
| | | | | | | | | | | | | | | |  
Db 2167 TTACGACAATAATTTTGGGAATATACCTCCTTAATAGGCCTAATTCGTCAATTGGGGA 2228  
| | | | | | | | | | | | | | | |  
QY 416 agcttgaaacattcacatctatagagctcaaaaaaacacatccaagaatactaactt 475



```
Db 2227 ATCTAACAACCTGTGCTAGCTGTGTATCTTACAAATATCAGCTTCTGCTCTATTCCG 2286
QY 476 CCGAacttggaatctgaagaatcctcatcagcttgatctgtacaacaacaatcttacag 535
Db 2287 CTTCAATTGGCAATATGAGAAATCTCAAACTCTGTTCTCAGTGAATACGATCTCATTG 2346
QY 536 ggaagttcccaacttcttgggaaaaatgaagtcctcgtgcttcttcaagcttaagaa 595
Db 2347 GGGAAATTCCTTCATTGTGTGCAATTTGACATCACTGGAAGTGTGTATATGTCAGAA 2406
QY 596 accgattgaccggtccaatccctagagcaactcaagcgaatcccaagccttaagtttg 655
Db 2407 ACAATTTGAAGGGAAGGTTCCGCAATGTTGGTAATATCAAGTGAACCTTCAATTTGT 2466
QY 656 acgtccaagaatgattgtgtgggacaatcccaaca 693
Db 2467 CGATGTCATCTAATAGTTTCAGAGAGAGAGCTCCCTCA 2504
```

Search completed: June 23, 2000, 22:52:36  
Job time: 40552 sec







OTHER INFORMATION: sativa)"  
US-08-475-891A-1

Query Match 5.3%; Score 58.6; DB 3; Length 6256;  
Best Local Similarity 50.8%; Pred. No. 7.6e-07;  
Matches 167; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

QY 365 tgaattgggaattcaaacctctctgacacatctgcgcgtgagcttggaagcttgaac 424  
DB 2735 TGGACATTGGGAGAAATACCTGGGGAGTCTTCTTAATGGTTCCAAATCTTCCA 2794  
QY 425 att---taacgatactagagctctcaaaaacacatcccaaggaactatcccgac 481  
DB 2795 CTTCGCTAGATTCTTTCGACATTGATTGAATTAAGATCAAGAGATTCCTCAAGATA 2854  
QY 482 ttggaatctgaagaatcctcagcttgagctgtacaaacaacatcttcaaggatag 541  
DB 2855 TTGGCAATCTATTTGGCTTACACATCTCTATCTCTGCAACAACAAATTTCAAGAGGTGAC 2914  
QY 542 ttccacattcttggaagaattgaagctctctgtcttcttaagccttaatgacacgat 601  
DB 2915 TTCCATCATCTGTTGGCAGGCTTAGAACTTAGCATTTAGTCCCTACGAAACACT 2974  
QY 602 tgaccggtccaatctcagagcaccacgcaatcccaagccttaagttgtgagct 661  
DB 2975 TGAGCGGTTCATCCCATGCGCATAGGAAATCTTACTGAACCTTAATCTTACTGCTCG 3034  
QY 662 caagcaatgattgtgtggacaatccca 690  
DB 3035 GCACCAACAATTCAGTGTGATACCA 3063

RESULT 2  
US-08-567-375-1  
; Sequence 1, Application US/08567375  
; Patent No. 5952485  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Yuang  
; APPLICANT: Song, Wen-Yuang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Confering  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/567.375  
; FILING DATE: 04-DEC-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004.645  
; FILING DATE: 29-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475.891  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373.375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 023070-058930  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6256 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1648..4383, 5178..5513)  
OTHER INFORMATION: /product= "RRK-F"  
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease  
OTHER INFORMATION: resistance gene RRK-F from rice  
OTHER INFORMATION: (Oryza sativa)"  
US-08-567-375-1

Query Match 5.3%; Score 58.6; DB 4; Length 6256;  
Best Local Similarity 50.8%; Pred. No. 7.6e-07;  
Matches 167; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

QY 365 tgaattgggaattcaaacctctctgacacatctgcgcgtgagcttggaagcttgaac 424  
DB 2735 TGGACATTGGGAGAAATACCTGGGGAGTCTTCTTAATGGTTCCAAATCTTCCA 2794  
QY 425 att---taacgatactagagctctcaaaaacacatcccaaggaactatcccgac 481  
DB 2795 CTTCGCTAGATTCTTTCGACATTGATTGAATTAAGATCAAGAGATTCCTCAAGAGATA 2854  
QY 482 ttggaatctgaagaatcctcagcttgagctgtacaaacaacatcttcaaggatag 541  
DB 2855 TTGGCAATCTATTTGGCAGGCTTAGAACTTAGCATTTAGTCCCTACGAAACACT 2974  
QY 542 ttccacattcttggaagaattgaagctctctgtcttcttaagccttaatgacacgat 601  
DB 2915 TTCCATCATCTGTTGGCAGGCTTAGAACTTAGCATTTAGTCCCTACGAAACACT 2974  
QY 602 tgaccggtccaatctcagagcaccacgcaatcccaagccttaagttgtgagct 661  
DB 2975 TGAGCGGTTCATCCCATGCGCATAGGAAATCTTACTGAACCTTAATCTTACTGCTCG 3034  
QY 662 caagcaatgattgtgtggacaatccca 690  
DB 3035 GCACCAACAATTCAGTGTGATACCA 3063

RESULT 3  
US-08-587-680A-1  
; Sequence 1, Application US/08587680A  
; Patent No. 5977434  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Yuang  
; APPLICANT: Song, Wen-Yuang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Confering  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004.645  
; FILING DATE: 29-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475.891  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373.375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774



```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
; OTHER INFORMATION: sativa)"
; US-08-587-680A-1

```

Query Match 5.38; Score 58.6; DB 4; Length 6256;  
Best Local Similarity 50.88; Pred. No. 7.6e-07;  
Matches 167; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

```

QY 365 tggatttggaatcaaacctctctgacaccttgcgcctgagcttgggaagcttgaac 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2735 TGGACTTGGAGAAATAAAGCTGGGGGGAGTCTTCCATAATGCTTTCCAACTTCCA 2794

QY 425 att---tacagatctagagctctacaacaaacacatccaaagacataccttcgaac 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2795 CTTCCTAGATTCTTCTCCACTTGATTGAAATAAGATCACAGGAGCATTCGAAAGATA 2854

QY 482 ttggaatctgagaatctcatcagcttggatctgttacaacaaacacatccaaagagtag 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2855 TTGGCAATCTTATTTGGCTTACAAATCTCTATCTCTGCAACAAATAATTGAGAGGTCAC 2914

QY 542 ttccacattctctgggaatgaagtcctcgtctcttcttaacggttaagacaccgat 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2915 TTCCATCATCTGTTGGGCGAGGCTTAGAAACTTAGCATTTCTAGTCCCTACGAAACAACT 2974

QY 602 tgaacggtccaaatctctagacactcagcggaatcccaagccttaagttgttgagctc 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2975 TTAGCGGTTCCATCCCATTTGGCCATAGGAAATCTTACTGAACTTATATCTTACTGCTCG 3034

QY 662 caagcaatgattctgttgggaacatccca 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3035 GCACCAACAAATTCAGTGTTGATACCA 3063

```

RESULT 4  
US-08-567-375-3  
; Sequence 3, Application us/08567375

```

; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Rorald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1..2676, 3520..3918)
; OTHER INFORMATION: /product= "Xa-21"
; US-08-567-375-3

```

Query Match 5.28; Score 57.4; DB 4; Length 3921;  
Best Local Similarity 48.18; Pred. No. 1.3e-06;  
Matches 163; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

```

QY 382 aacctcttgacactcttgcgcctgagcttgggaagcttgaacatttaacgtatcagag 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1159 AAGATCAGAGAACATTCGGAAGATATGCGAATCTTATTTGGCTTACAAATCTCTAT 1218

QY 442 ctctcaaaaacacatccaaaggaatataccttcggaacttggaaatctgaaagatc 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1219 CTCTCAACAACAATTCAGAGGCTCTCTCCATCATCTGCGGCGAGGCTTAATAAACTTA 1278

QY 502 atcagcttgatctgttacaacaaacatctacagagatagttccacttcttgggaana 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1279 GGCAATTTACTCGGCTACGAGAAACAATTCAGAGGCTTCGATCCCTGTTGGCCATAGGAAAT 1338

QY 562 ttgaagctctggtcttcttaacggttaatgaacacgattgacgcggtccaatcctagag 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



Db 1339 CTTACTGAACCTTAATATCTTACTGCTGCGACCAACAATATCAAGTGGTATGATACCATAC 1398  
QY 622 cactacggaatcccaagccttaagttgtgacgtcacaagcaatgattgtgtg 681  
Db 1399 ACATCTCAAAACCTCACAACCTTGTGTCAATTAGCCTTCACTAATAACCTTAGTGT 1458  
QY 682 acaatccacaacagcagcccttgctcacatccctta 720  
Db 1459 CCAATACCCAGTGAATATTATTCATATTCAAACACTATCA 1497

## RESULT 5

US-08-587-680A-3  
; Sequence 3, Application US/08587680A  
; Patent No. 5977434  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Conferring  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/587,680A  
; FILING DATE: 17-JAN-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475,891  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004,645  
; FILING DATE: 29-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/567,375  
; FILING DATE: 04-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 023070-058940US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3921 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1..2676, 3520..3918)  
; OTHER INFORMATION: /product="Xa-21"  
US-08-587-680A-3

Query Match 5.2%; Score 57.4; DB 4; Length 3921;

Best Local Similarity 48.1%; Pred. No. 1.3e-06;  
Matches 163; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 382 aacctctcgcacatcttcgcctgaagcttgggaagcttgaacattacagatcagag 441  
Db 1159 AAGATCAGCAGGAAGATCCGAGAGATATGCAATCTTATGCTTACAACTCTAT 1218  
QY 442 ccttcaaaaataacatccagaagactatccttcgcgaacttgaatctgaagaactc 501  
Db 1219 CTTGCAACAAATTTAGAGGGGTCTCTTCATCATCTGTTGGCAGGCTTAAAACTTA 1278  
QY 502 atcagcttgatctgtacaacacatcttacaggatagttccacttctctggaaa 561  
Db 1279 GGCATTCTACTGCTGCTAGAAAACAACTTGAGCGGTTCGATCCGTTGCCATAGAAAT 1338  
QY 562 ttgaagctcgtgtcttcttaccgcttaatgaacacgattgacgcgtccaatcctgag 621  
Db 1339 CTTACTGAACCTTAATATCTTACTGCTGCGACCAACAATTCAGTGTGATGCATAC 1398  
QY 622 cactacggaatcccaagccttaagttgtgacgtcacaagcaatgattgtgtg 681  
Db 1399 ACATCTCAAAACCTCACAACCTTGTGTCAATTAGCCTTCACTAATAACCTTAGTGT 1458  
QY 682 acaatccacaacagcagcccttgctcacatccctta 720  
Db 1459 CCAATACCCAGTGAATATTATTCATATTCAAACACTATCA 1497

## RESULT 6

US-08-475-891A-3  
; Sequence 3, Application US/08475891A  
; Patent No. 5859339  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; TITLE OF INVENTION: Procedures and Materials for Conferring  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,891A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; PRIOR APPLICATION DATA:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 023070-058910US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5992 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:



```

: NAME/KEY: CDS
: LOCATION: join(512..3149, 3993..4393)
: OTHER INFORMATION: /product= "RRK-B"
: OTHER INFORMATION: /note= "xal xanthomonas spp. disease
: OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
: OTHER INFORMATION: sativa)"
: JS-08-4/5-891A-3

```

US-08-475-891A-3

Query Match	5.2%;	Score 57.4;	DB 3;	Length 5992;
Best Local Similarity	48.1%;	Pred. No. 1.6e-06;		
Matches 163; Conservative	0;	Mismatches 176;	Indels 0;	Gaps 0;

OY	382	aacctctctggaactctctggccctggagctctggaaacttgaaactttacagatctag	441
Db	1670	AAAGTACAGGAAGACATTCGGAGGATATTGGCAATCTTATGGCTTACAACTCTCAT	1729
OY	442	ctctacaaaaacaacatccaaaggaactatacctccgaacttggaaatctgaagaatc	501
Db	1730	CTTCGCACACACAAATTTCAAGAGGCTCTCTCCATCTATCTTGGCAGGGTTAAAACTTA	1789
OY	502	atcagcttggatctgttacaacaacaactcttaagsgatagttcccaacttcttggaa	561
Db	1790	GGCATTTCTACTCGGCTTCGAAAAACACTTtaggggtttgcattcccttggccatagaaat	1849
OY	562	ttagaagtctctgctcttctttaaagcttaatgacacgcatgacggctccaatcc	621
Db	1850	CTTACTGAACTTAATATCTTACTGCTCTGCGACACMAAATTAAGTGTGGATACCAATC	1909
OY	622	cactcaagcgcatcccaagctctaaagtgtgtgacgtctcaagcaaatgattgtgg	681
Db	1910	ACATCTCAAAACCTCAACAACCTGTGTGCTATTAGGCTTCAACATTAACCTTAGTGT	1969
OY	682	acaatcccaacaagsgaaccttgtctacattccctta	720
Db	1970	CGAATACCCAGTGAATTTATTCATATTTCACAACTATCA	2008

RESULT 7  
PC-08-587-

; Sequence 24, Application US/08587680A

; Patent No. 5977434

? GENERAL INFORMATION:  
? APPLICANT: Ronald, Pamela C.  
? APPLICANT: Wang, Guo-liang  
? APPLICANT: Song, Wen-Yuang  
? APPLICANT: Szabo, Veronique  
? TITLE OF INVENTION: Procedures and Materials for Confering  
? TITLE OF INVENTION: Disease Resistance in Plants  
? NUMBER OF SEQUENCES: 27  
? CORRESPONDENCE ADDRESS:  
?

ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:

```

:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/587,680  
 FILING DATE: 17-JAN-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/373,375  
 : FILING DATE: 17-JAN-1995  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/475,891  
 : FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 60/004,645  
;; FILING DATE: 29-SEP-1995

;  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/567,375

.. FILING DATE: 04-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Bastian Kevin T.

PROJECT STATION NUMBER: 34

REGISTRATION NUMBER: 34114

REFERENCE/DOCKET NUMBER: 023070-05894005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEO ID NO: 24

## SENTENCE CHARACTERISTICS

LENGTH: 1554 base pairs

WENTON, JOSE WESLEY  
MYND, JAMES WESLEY

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: Linear

MOLECULE TYPE  
US-08-587-680A-24

Query Match	5.1%;	Score 56.4;	DB 4;	Length 1554;
Best Local Similarity	49.0%;	Pred. No. 1.6e-06;		
Matches 150; conservative	0;	Mismatches 156;	Indels 0;	Gaps 0;

QY	384	ccctctggacacatcttgccgcctgaagcttggaaacttgacattcacatcttagagct	44
Db	18	CTTGTCGGGACACTTCCTAGTGTCTATTGGAACTATTCCAGGCTCAAAACCTTGTTT	77
QY	444	ctacaaaaaanaaatccaaggaactataccttcggaacttggaaatctgaagaatcat	503
Db	78	AACGTGAAGTGTCTCTCAGGAGTATCCCTTCGATTTGGCAGACTAAAGCATCTT	137
QY	504	cagcttgatctgtatacaacaacacttaagaagatagttcaccactctcttggaaatt	565
Db	138	AAAGCTGGACCTCAATGAAAGAAACTTCTCGGACAAATCCCTCCATATTTGGTAATG	197
QY	564	gaagctctctgtctctttagcagctaatagacaaccgattgaccggttcacatcttagaca	623
Db	198	TCCTTCCTTAACCTTAACCTTGATTTAGCCAAATCAACTTTCGGTCTATCCCACTTCA	257
QY	624	ctcagcggacaaccacaagcctttaagttgttgaagcttcagaagaatgatgttgtggagc	683
Db	258	AATGTGCATTAATTCACATCTTAATTAATCATCAATATTTCCGTGAATCACTTCAAGAGAG	317
QY	684	aatccc	689
Db	318	CCCTTCC	323

## RESULT 8

US-08-238-163-3  
: Sequence 3 Application MS/08238163

Patent No. 5,569,830

```

1  GENERAL INFORMATION:
2  APPLICANT: BENNETT, Alan
3  APPLICANT: LABAVITCH, John M.
4  APPLICANT: POWELL, Ann
5  APPLICANT: STOLT, Henrik
6  TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
7  TITLE OF INVENTION: POLYCALCITRONASSS AND THEIR USE TO CONTROL FUNGAL DISEASE
8  NUMBER OF SEQUENCES: 24
9  CORRESPONDENCE ADDRESS:

```

CONTRACTOR: Townsend and Townsend Kourie and Crew  
ADDRESS: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California

```

;      COUNTRY:  US
;      ZIP:      94105-1493
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;

```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,163
; FILING DATE: 03-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 421..1401
; US-08-238-163-3

```

```

Query Match      4.9%; Score 54.4; DB 1; Length 2075;
Best Local Similarity 50.3%; Pred. No. 6.7e-06;
Matches 190; Conservative 0; Mismatches 181; Indels 7; Gaps 2;

```

```

QY 242 tctacgctctcgcgaggtttacagatccagaccatgtccctccagaagctggatccaa 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 513 TCTACAAATTAAGAAGACTTAGGCAATCCTTACATTAGCTTATGGAGATCCAAACAC 572
QY 302 ctcttgtaacctgttacctgttccatgctcaacccgttaacacacccgctcacc 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 573 AGATTGCTGTACTGTACTGACTATAAATGTGACCGGAAAC---CAACCGATTAAAG 628
QY 362 gtgtgatttggaagaattcaacctctctgacatcttgagccttgagctgggaagcttg 421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 629 CTCTACACGCTTCCAGACCAATATCTCCGGCCAATTCGGGACCGCTCGGAGACTTC 688
QY 422 aacattacagatctagagctctacaa---aaacacatccaaaggaaactatccttcg 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 689 CATATCTCGAAGATTCGAATTCATCATCTACTATCATCCAGGACCAATTCACACG 748
QY 479 aacttgaagaatctgaagaatctatcagcttgatctgttcaacacacatcttaaggga 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 749 CAATTGGGAAGCTCACAAATCTCAAAATGTTAAGGCTCAGCTTCACCTTAACCTTACAGGTC 808
QY 539 taattccaccttcttggaagaattgaagctctgtgtctttttagcgcttaataagacaac 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 809 CGATCCCTGATTCCTTAGTCAGCTGAAGATTGACGTTGCTCGAGTTGATTAATACATC 868
QY 599 gattgaccggtccaatcc 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 869 AATTACCGGAACAAATCC 886

```

```

RESULT 9
US-08-238-163-1
; Sequence 1, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAYTCH, John M.
; APPLICANT: POEHL, Ann
; APPLICANT: STORZ, Henrik
; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
; TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew

```

```

; STREET: Stewart Street Tower, One Market Plaza
; City: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,163
; FILING DATE: 03-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1058 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..1023
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1058
; OTHER INFORMATION: /standard_name= "Pear PGIP cDNA"
; US-08-238-163-1

```

```

Query Match      4.8%; Score 53; DB 1; Length 1058;
Best Local Similarity 51.5%; Pred. No. 1.2e-05;
Matches 122; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

```

QY 380 caaacctctctggaacatcttgcgcttagcttgggaagcttgaacattacagatctag 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 347 CCAATCTCACTGCGCCATCCAAACCCCAATGCGCAAGCTCAAGACTCAAGTCTCTCA 406
QY 440 agcttacaagaagaatccaaagaaacttaccttcgaaacttgaatctgaagatc 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 407 GCTTAGCTGACCAACTCTCAAGCTCTGTCTCTGACTTCTTCAAGCACTCAAGAAC 466
QY 500 tcatcagcttgatctgtacacacaaacttatacaggatagttccacttcttggaa 559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 467 TCACATCTCTGAGACTCTCTTCAACAACCTCACCGGAGCCATCCAGCTCGCTTCTG 526
QY 560 aattgaagctctcggtctttttagcgcttaataagacacagattgacggtccaatcc 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 527 AGCTCCAAACCTCGCGCTCTTGTCTAGACCGCAATAAAGCTCACAGGTGATATTTC 583

```

```

RESULT 10
US-08-567-375-15
; Sequence 15, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

```











Tue Jun 27 14:48:03 2000

```

1  PRIOR APPLICATION DATA:
2  APPLICATION NUMBER: WO PCT/IT/00158
3  FILING DATE: 04-DEC-1992
4  ATTORNEY/AGENT INFORMATION:
5  NAME: Sullivan, Sally A.
6  REGISTRATION NUMBER: 32,064
7  REFERENCE/DOCKET NUMBER: 19-94
8  TELECOMMUNICATION INFORMATION:
9  TELEPHONE: (303)499-8080
10 TELEFAX: (303)499-8089
11 INFORMATION FOR SEQ ID NO: 16:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 792 base pairs
14 TYPE: nucleic acid
15 STRANDEDNESS: double
16 TOPOLOGY: linear
17 MOLECULE TYPE: cDNA
18 HYPOTHEICAL: NO
19 ANTI-SENSE: NO
20 ORIGINAL SOURCE:
21 ORGANISM: Phaseolus vulgaris
22 STRAIN: Pinto
23 IMMEDIATE SOURCE:
24 CLONE: pMD-1 (DSM NO. 57446926821)
25 FEATURE:
26 NAME/KEY: CDS
27 LOCATION: 2..685
28 FEATURE:
29 NAME/KEY: 3'UTR
30 LOCATION: 686..792
31
32 US-08-244-646-16

```

```

Query Match      3.7% ; Score 40.8 ; DB 2; Length 792;
Best Local Similarity 48.0% ; Pred. NO. 0.025;
Matches 117; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY    382 aaectctctggacatcttgcgacctgaagcctttggaacttgtaacatttcagtlactagag 441
      |||||
Db     5 AACCTCCTCGTGCCATACCCCCCGGCATCGGTAAACACCACAATCCACTATCTCAT 64
      |||||

QY    442 ctccacaanaaacacatccaaaggaaactataccttcgcgaacttggaaatctgaagaatctc 501
      |||||
Db     65 ATCACCCCAKACCAAATGTCCTCCGGCCGAATCCCAGATTTCCTTGTCAKAGATCAAAAACCTC 124
      |||||

QY    502 atcacgttgtatctgtfacacaacaacatcttacaggatagttcccaattcttggysaaa 561
      |||||
Db     125 GTCAACCTCGACTTCTCTCTACAAAGGCCCTCTCCGGCACCCATACCTCCCTCATCTCTCT 184
      |||||

QY    562 ttgaagctcttggtctttttaaggcttaagtgaacacgatggaccggtccaatcctagag 621
      |||||
Db     185 CTCGCCAACCTGTAGAATAACATTCAGACGGCAACCAATCTCGGGCGCCATCCCGGAC 244
      |||||

QY    622 caact 625
      ||
Db     245 TCCCT 248

RESULT 15
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Meckiewicz & No. 5389526rls
STREET: One Liberty Place 46th Floor
CITY: Philadelphia

```

```

1 STATE: PA
2 COUNTRY: USA
3 ZIP: 19103
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: PatentIn Release #1.0, Version #1.25
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/07/867,106
13 FILING DATE: 19920625
14
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: AU PJ 7187
17 APPLICATION NUMBER: PCT/AT90/00530
18 FILING DATE: 02-NOV-1989
19
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Feeney, Joanne Longo
22 REGISTRATION NUMBER: 35,134
23 REFERENCE/DOCKET NUMBER: RICE-0002
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: 215-568-3100
26 TELEFAX: 215-568-3439
27
28 INFORMATION FOR SEQ ID NO: 2:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 5852 base pairs
31 TYPE: NUCLEIC ACID
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: DNA (genomic)
35 ANTI-SENSE: NO
36
37 FEATURE:
38 NAME/KEY: CDS
39 LOCATION: 2378..5038
40
41 FEATURE:
42 NAME/KEY: CDS
43 LOCATION: 2378..5038
44
45 US-07-867-106-2

```

	Query Match	3.6%	Score 40;	DB 1;	Length 5852;
	Best Local Similarity	47.3%	Pred. No. 0.1;		
	Matches 121;	Conservative	0;	Mismatches 135;	Indels 0; Gaps 0;
QY	848	aacaccctcacccattatccaatacaccatcatbalytaalaagtatatatatytagia	907		
Db	5680	AATATTAAATTTTAAATTAATAAATAAAAAAAATTTTTAAATTTAAAAAAA	5621		
QY	908	aaacaacaaaaaagaagatcgatcgtaactgatcatcgtgcacattgagaacttcg	967		
Db	5620	AAAAAAAAAAAAAAAAAAAAAAAAAATAAAATGTAATTTTGGAATTAATAAATTAATAATGTA	5561		
QY	968	aggctcgtatgtaaaatttctaargcaatttcggcttaactgtaaytttcggitgigga	1027		
Db	5560	AAGGGTTTTTTTTTAAATATGTACGATTTTTTTTAAATCATTTGAGCAATTTAAAA	5501		
QY	1028	ttcgcgaagctaacattgtaacttgtaigtatagaatcagttgctcgttcgttcgcaaaa	1087		
Db	5500	AATCTTAAACATTAACAATATTGATTTTTTTTTTTTTTTTTTTTTTTTAAAAAAT	5441		
QY	1088	aaaaaaaaaaaaaaaaa	1103		
Db	5440	TCAAATTAATTAATAA	5425		

Search completed: June 23, 2000, 22:44:52  
Job time: 40238 sec



7  
.  
.  
.

.....











REFERENCE  
1 (bases 1 to 447)  
AUTHORS  
Newman,T., deBruin,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raiknel,N., Somerville,S., Thomasow,M.,  
Retzel,E. and Somerville,C.  
TITLE  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL  
Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE  
95148729  
COMMENT  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcne@dm.cl.msu.edu  
Seq primer: T7 dye primer.

FEATURES  
SOURCE  
Location/Qualifiers  
1..447  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="179A227"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: lambda Zip-Lox; Site.1: Sal; Site.2: Not;  
Lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques. The vector is BRU's lambda Zip-Lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA."

BASE COUNT 106 a 115 c 83 g 126 t 17 others  
ORIGIN  
1..556  
Location/Qualifiers

Query Match 34.2%; Score 378.4; DB 23; Length 447;  
Best Local Similarity 93.1%; Pred. No. 2.4e-78;  
Matches 418; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

17 tccaactcaataaagggaacacagtaaccctaatttgccttctccttctgta 76  
1 tccacattcaataaaggggg-aaaccaactaaccttaatttttctccttctgta 59  
77 gaaatttcccttaccctcaaatctcttgcattcctcctctaaccgcgaag 136  
60 gaaatttcccttaccctcaaatctcttgcattcctcctctaaccgcgaag 119  
137 ctcaatgagctcgaacactacggtgagctcttcgacgcttgtaaccctaact 196  
120 ctcaatgagctcgaacactacggtgagctcttcgacgcttgtaaccctaact 179  
197 tagcttattcaactggtcgaacacactcgaagagatgctcttaactccttcgc 256  
180 tagcttattcaactggtcgaacacactcgaagagatgctcttaactccttcgc 238  
257 ggaattgacagatcacaacatgctctcagagctggatccaactctgttaactc 316  
239 ggaattgacagatcacaacatgctctcagagctggatccaactctgttaactc 298  
317 gfacctggttcaatgctcctgtaacgaacacacgcttactcgtgtgattggaa 376  
299 gtaacgggttcaatgctcctgtaacgaacacacgcttactcgtgtgattggaa 358  
377 atcaaacctctctggaacatctgc-gcctagcttggaaacttgaaacttaacgat 435  
359 attttaaactcttctggaacatctgc-gcctagcttggaaacttgaaacttaacgat 418  
436 cttagagctctaaacaaacatccaag 464

Db 419 NTAGAGNTTCAAAAACAACATCAGGG 447

RESULT 3  
LOCUS  
DEFINITION  
R89998 556 bp mRNA EST 30-DEC-1997  
16353 Lambda-PRL2 Arabidopsis thaliana cDNA clone 187H577, mRNA  
sequence.  
ACCESSION  
R89998  
VERSION  
R89998.1 GI:957538  
KEYWORDS  
EST.  
SOURCE  
thale cress.  
ORGANISM  
Arabidopsis thaliana

REFERENCE  
1 (bases 1 to 556)  
AUTHORS  
Newman,T., deBruin,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raiknel,N., Somerville,S., Thomasow,M.,  
Retzel,E. and Somerville,C.  
TITLE  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL  
Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE  
95148729

COMMENT  
On Apr 14, 1993 this sequence version replaced gi:693023.

CONTACT: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcne@dm.cl.msu.edu  
Seq primer: T7 dye primer.  
Location/Qualifiers  
1..556  
Location/Qualifiers

FEATURES  
SOURCE  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="187H577"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: lambda Zip-Lox; Site.1: Sal; Site.2: Not;  
Lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques. The vector is BRU's lambda Zip-Lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA."

BASE COUNT 151 a 118 c 123 g 146 t 18 others  
ORIGIN  
1..556  
Location/Qualifiers

Query Match 32.4%; Score 358; DB 23; Length 556;  
Best Local Similarity 92.1%; Pred. No. 1.4e-73;  
Matches 431; Conservative 0; Mismatches 29; Indels 8; Gaps 5;  
385 ctctctggaactcttggtcgtgagcttgggaacattggaacatttaagatagatgc 444  
1 ctctctggaactcttggtcgtgagcttgggaacattggaacatttaagatagatgc 60  
445 tacaanaaacatccaaggaactataccttcgaacttgaacttgaactgaatctcatc 504  
61 tacaanaaacatccaaggaactataccttcgaacttgaacttgaactgaatctcatc 120  
505 agcttgagcttgtaacaaacaaacttaacaggaatggtcccaacttcttggaatg 564  
121 agcttgagcttgtaacaaacaaacttaacaggaatggtcccaacttcttggaatg 180  
565 aagctctggtcttcttaagcgttaatgaacacgattgaccggttccaat-cttagaaga 623



|||||  
 Db 181 AAGTCTGTGCTTTTACGGCTTAATGACACCGATTGACGGCGCCCAATCCCTAGAGCA 240  
 QY 624 ctaaggcaatcccaagccttaagattgttcgctctaagaatgattgttgagac 683  
 Db 241 CTCACGTGCAATCCCAAGCC-TTAAAGTGTGATGTCTCAAGCAAGATTTGTGTGAC 299  
 QY 684 aatccacaacaagcagcccttgcacatcccttaccagaacttgaagaacaccgag 743  
 Db 300 AATCCCAACAAGGAGACTTGTCTGACATCTCTTACAGAACTTGAAGACACCCGAG 359  
 QY 744 attgg-agggaccggaattcctgcttcgcaagcgaacactaactgcaactgaa 802  
 Db 360 GTTGGNAGGAGCCGGAATTACTGGTCTTGCAGCTNCGCAATTAACGAC--TGAA 416  
 QY 803 actggcaaacctgaatgaagaattgggggtgaacctgtgaagac 850  
 Db 417 AATTGAAAACCTGAAA--TGAGATTTGGGGGTGNCCTTAAAGCAC 462

RESULT 4  
 AA712221 401 bp mRNA EST 24-DEC-1997  
 LOCUS 31949 lambda-PRL2 Arabidopsis thaliana cDNA clone 180C977, mRNA  
 DEFINITION  
 ACCESSION AA712221  
 VERSION AA712221.1 GI:2722138  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.  
 1 (bases 1 to 401)  
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
 McIntosh,L., Ohlroge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 On Sep 19, 1997 this sequence version replaced gi:1520373.  
 Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313cne@lhm.cl.msu.edu  
 Seq primer: T7 dye primer.  
 Location/Qualifiers  
 1..401

## FEATURES

Location/Qualifiers  
 1..401  
 /organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="180C977"  
 /clone\_id="lambda-PRL2"  
 /note="Vector: lambda zip-lox. Site\_1: Sal. Site\_2: Not;  
 lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRU's lambda zip-lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 c1190 df primed cDNA. "

BASE COUNT 103 a 105 c 78 g 99 t 16 others  
 ORIGIN

Query Match 31.1%; Score 343.6; DB 37; Length 401;  
 Best Local Similarity 92.1%; Pred. No. 3.1e-70;  
 Matches 363; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 121 cttaaacctcgaagctcacatgagcgtctcgaaactatcggtggaagctcttcgaagct 180  
 Db 1 CTTAAACCCGAAAGCTCATGAGCTCTCGAAACTATCGGTGGAGCTCTTCGAGCT 60  
 QY 181 tcgttaacccttaactagctcttgatcacctggtgcgaagcaactcgaagaagatgct 240  
 Db 61 TCGTTAACCCCAACCTTAGCTTATTCACCGNGTCGAAGCAACNCCGAGNANNGCT 120  
 QY 241 cttcaagcctctggccggaattggaagatccagacatgctcctcgaagctggaatca 300  
 Db 121 CTCTACGCTCTCCGCGAGTTTGACAGATCCAGACCATGCTCTCCAGAGCTGGATCCA 180  
 QY 301 actctgttaacctgttacctgttcacatgtaacctgttaaccacaagaacacccgctact 360  
 Db 181 ACTMTGTNNATCCTTGTACNGSTTCCATGTNACCTGTNACCAAGCAACCGCTACT 240  
 QY 361 cgtgtggaattgggaattcaaacctctcgtgacacctctgagcctgtaagcttggaagct 420  
 Db 241 CGTGTGATTTGGGAAATTCAAACNNCTCTGACATCTTCCGNTAGCTTGGGAAAGCTT 300  
 QY 421 gaacattacagatctacagcctcacaacaacaacacacccaagaactatccttcgaa 480  
 Db 301 GAACATTACAGTATCTTAGACTCTTACAAAACCAACCAACCAAGCAAGCACTATACNTCCGA 360  
 QY 481 ctgggaactctgaagaatctcatcagcttgagct 514  
 Db 361 CTGGAATCTGAAG-ATCTCATCAGNTTGGATC 393

RESULT 5  
 AA394359 534 bp mRNA EST 30-OCT-1997  
 LOCUS 25942 lambda-PRL2 Arabidopsis thaliana cDNA clone 305G177 3, mRNA  
 DEFINITION  
 ACCESSION AA394359  
 VERSION AA394359.1 GI:2047570  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.  
 1 (bases 1 to 534)  
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
 McIntosh,L., Ohlroge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 On May 18, 1995 this sequence version replaced gi:811121.  
 Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313cne@lhm.cl.msu.edu  
 Seq primer: T7.  
 Location/Qualifiers  
 1..534

## FEATURES

Location/Qualifiers  
 1..534  
 /organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="305G177"  
 /clone\_id="lambda-PRL2"  
 /note="Vector: lambda zip-lox. Site\_1: Sal. Site\_2: Not;

## FEATURES

Location/Qualifiers  
 1..534  
 /organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="305G177"  
 /clone\_id="lambda-PRL2"  
 /note="Vector: lambda zip-lox. Site\_1: Sal. Site\_2: Not;







SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 442)  
AUTHORS Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.  
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE 95148729  
COMMENT On Jan 14, 1998 this sequence version replaced gl:1797474.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313c@pml.msu.edu  
The sequence entry for this EST has been reverse complemented and is being submitted in the sense orientation.  
Seq primer: M13 Universal.

FEATURES  
source Location/Qualifiers  
1..442  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="90J7XP"  
/note="Vector: lambda ZAP-Lox. Site\_1: Sal. Site\_2: Not; lambda PR2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda ZAP-Lox. The cDNA inserts were directionally cloned with Sal-I Not arms using 0.1µg dt primed cDNA."

BASE COUNT 135 a 82 c 90 g 132 t 3 others  
ORIGIN

Query Match 29.6%; Score 327.8; DB 42; Length 442;  
Best Local Similarity 89.4%; Pred. No. 1.5e-66;  
Matches 397; Conservative 0; Mismatches 40; Indels 7; Gaps 4;

QY 644 ttaagttgtgacgtcaagcaatgtgtgtggacaat-cccaacaacggacc 702  
DB 1 TAAAGTTTGTGATCTTCACACCATNGTTGTGGAACAATCCCAACAAACGACT 60  
QY 703 ttgtgcacatctcttaccagaacttggagacaacccggagattgg--agggaacggaa 760  
DB 61 TTGCTCCACATTCCTTTACAGACTTTGAGAACACCCCGAGGTGGAGCGACCGAAT 120  
QY 761 tactggtcttgcagctcagacactgaactgaacacactgggaaacactggaaa 820  
DB 121 TACTGCTTCGACGCTACGACACTGACCTGGAATAATGGCAAAACCTGAA 180  
QY 821 tgaagaatcggggggtgacacttgaagaacacttccacttacaatacatcatca 880  
DB 181 TGAAGAATGGGGGGGTGACCTGTAGAACACTTCACCACTTATCAATATACATCTA 240  
QY 881 ttatgtataatgatatataatgtaaaaaaataaagaagaatcgatcgtaatt 940  
DB 241 CTATGTAAATGATATATATATGTAGTCAAA--AAAAAATGAAATGATCGATCTAT 298  
QY 941 atcatcgtctcaatcgaacttcgaggtc--tgatgtataaatttcaatcgatt 998  
|||||

DB 299 ATCACTGTGCTCAATTCGAGAATTTGAGGCTGTGTATGTAATAATTTCAATCGACT 358  
QY 999 ttgcctactgtaatgttcggttggagatcagaagaataacttgtattgtatgt 1058  
DB 359 TTCGGTACTGTAAATGCTGGTGGGATCTCGAAGTAACTTTGTATGTGTATGCT 418  
QY 1059 atcaatgtctgcctgtctgc 1082  
DB 419 ATCAAGTTTCTGCTGTGCTGC 442  
|||||

RESULT 8  
234187/c 336 bp mRNA EST 14-JUN-1994  
LOCUS ATMS3221 Versailles-VB Arabidopsis thaliana cDNA clone VBVDH02,  
DEFINITION mRNA sequence.  
ACCESSION 234187  
VERSION 234187.1 GI:498544  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 336)  
AUTHORS CNRS.  
TITLE The Arabidopsis thaliana transcribed genome: the GDR cDNA program  
JOURNAL Unpublished (1996)  
COMMENT Contact: Desprez T., Amselem J., Chapello H., Rouze P., Caboche M., Hofte H.  
INRA Versailles  
Laboratoire de Biologie Cellulaire  
Route de Saint-Cyr 78026 Versailles Cedex, France  
Email: thierry@versailles.inra.fr.

FEATURES  
source Location/Qualifiers  
1..336  
/organism="Arabidopsis thaliana"  
/strain="ecotype Columbia"  
/db\_xref="taxon:3702"  
/clone="VBVDH02"  
/clone\_lib="Versailles-VB"  
/tissue\_type="whole seedlings"  
/dev\_stage="in vitro-grown etiolated seedlings, 5 days old"  
/note="Vector: pBluescript"

BASE COUNT 89 a 62 c 65 g 119 t 1 others  
ORIGIN

Query Match 29.6%; Score 327.2; DB 20; Length 336;  
Best Local Similarity 98.8%; Pred. No. 2e-66;  
Matches 329; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 671 atttggtgtggaatcccaacaacggacccctgtgcacacttcttaagaacttg 730  
DB 333 ATTTGTGTGGAACAACTCCCAACAACGACCTTGTCTCAGCAATCTTACGAACCTTG 274  
QY 731 agacaacccgagattggagggacggaaattactcgtcttgcagctcagacactaact 790  
DB 273 AGAACAACCGAGATGGAGGGACGGAAATTAAGTCTTCAAGTACGACATAACT 214  
QY 791 gcaactgaacaacactggaacactgaaatgaagaatctgggggtgacacttgaagaac 850  
DB 213 GCACCTGGAACAACCTGCAAAACCTGAATAAGAAATGGGGGCTGACTGTGAAGAAC 154  
QY 851 acttcacacttataataatcacatctatgttaataagtatataatgtatga 910  
DB 153 ACTTCACCACTTTATCAATAATACATCTATATATGTAATAATATATATAGTACTAAA 94  
QY 911 acaaaaaaatagaagaatcgtaactcgtaatacatcctgtgtcctcaattgaagaactcgag 970  
DB 93 ACAAAAAAATGAAGATGAATCGGTATATCACTGTGTCAATGTGAAGAACTCGAGG 34  
|||||



```

QY 971 tctgtatgtaaaattcttaaatcgatttcgc 1003
Db 33 TCTGTATGTAATAATTCTTAAGTCGATTTCGC 1

RESULT 9
AL100685 360 bp mRNA EST 21-AUG-1998
DEFINITION 31806 Lambda-PRL2 Arabidopsis thaliana cDNA clone 90J7XP 3', mRNA
sequence.
ACCESSION AL100685
VERSION AL100685.1 GI:3449395
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsi.
REFERENCE 1 (bases 1 to 360)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
GENES galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL 95148729
MEDLINE
COMMENT On May 5, 1995 this sequence version replaced gi:797715.
CONTACT: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@dm.cl.msu.edu
The sequence entry for this EST has been reverse complemented and
is being submitted in the sense orientation.
Seq primer: M13-21.
FEATURES
SOURCE location/Qualifiers
1..360
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone_lib="90J7XP"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT 116 a 60 c 77 g 105 t 2 others
ORIGIN
Query Match 28.6%; Score 316; DB 42; Length 360;
Best Local Similarity 95.6%; Pred. No. 8,3e-64;
Matches 346; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
QY 723 gaacttgaagaacaaccgagattgagagacggaattctcggtctgcaagctacga 782
Db 1 GAACCTTGAAGAACACCCGAGGTTGGAGGACCGAATTACTGGTCTTCGAAGCTACCA 60
QY 763 cactactgcacctgaacacctggcaaaacctgaataaattgaggggtgacctt 842
Db 61 CACTTACTGCACCTGCAAAAATTGGCAAAACCTGAATAATGAGAAATTGGGGGTGACCTT 120
QY 843 gtaagaacacttcaccacttatacaaatatcatctattatgtaataagtatatatg 902

```

```

Db 121 GTAGACACCTTCACACCTTATCAAAATATCACACTCTATGTAATATATATATG 180
QY 903 tagtaaaacaaanaaataagaatcgatcgatgaatcatcatctgctcctaattgaa 962
Db 181 TAGTCCAA--AAAAAAGAAAGAAATCGAATCAGTAATATATCTGAGTCAATGAGAA 238
QY 963 cttagagatc--tgatgtaaaattcttaaatcgatttcgactgtaagtgcgt 1020
Db 239 CTTTGAGGTCGTGATGTAATAATTTCTAAATCGACCTTTCGCCACTGTAATGTCGGT 298
QY 1021 tctgagattctgagaagatcatctgtattgtagatgataaattgttcgcttctc 1080
Db 299 TGTGGATNCTGAGAAAGTAACATTTGTATGTGATGATCAAGTTGCTTCGCTGTCT 358
QY 1081 gc 1082
Db 359 GC 360

RESULT 10
H37300 353 bp mRNA EST 30-DEC-1997
LOCUS H37300
DEFINITION 15429 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179J19T7, mRNA
sequence.
ACCESSION H37300
VERSION H37300.1 GI:906799
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsi.
REFERENCE 1 (bases 1 to 353)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
GENES galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL 95148729
MEDLINE
COMMENT On May 8, 1995 this sequence version replaced gi:801146.
CONTACT: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@dm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
SOURCE location/Qualifiers
1..353
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone_lib="179J19T7"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT 98 a 85 c 67 g 89 t 14 others
ORIGIN

```



Query Match 28.0%; Score 309.6; DB 23; Length 353;  
 Best Local Similarity 94.4%; Pred. No. 2.6e-62;  
 Matches 334; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

445 tacaataacatccagaactatactcctcgaacttggaatctgaagaatctcacc 504  
 1 TACAAAACACATCCAGAGACTATACCTTCGAACTTGGAAATCTGAAGATCTCAAC 60

505 agcttgatctgtacaacaactctacaggaatgtccactctcttggaatattg 564  
 61 AGCTTGATCTGTACACAACTTACAGGATGATCCACTTTTGGGAAATG 120

565 aagctctgtcttttaagagcttaagacacccgattgacgggtccaat-cctagagca 623  
 121 AAGTCTGTGCTTTTACGGCTTAATGACACCGATTGACCGGTCCATCCCTAAGCA 180

624 ctacagcaatcccaagccttaaaagtgtgacgctcacaagaatgattgtgagac 683  
 181 CTCACGGCAATCCCAAG-CATTAAAGTTGTGANGTCTCANGCAATGATTTGTGGAAAC 239

684 aatcccaacaaaggacccttctgacatctccttcaagaatttgagagaaacccgag 743  
 240 AATCCCAACAAAGCGNCCCTTGTCTCAGTCCCTTACAGAACTTGGNAAACCCGNG 299

744 atggagagaccggaattactcgtctgtgcaagctacgacactgaactgacctg 797  
 300 ATTGAGAGGNCGGANTTACTCGTCTTGCAGATGACGCACTACATGACNGC 353

RESULT 11  
 A1100683 364 bp mRNA EST 21-AUG-1998  
 LOCUS 33804 Lambda-PRL2 Arabidopsis thaliana cDNA clone 187H5XP 3', mRNA  
 DEFINITION  
 sequence.  
 ACCESSION A1100683 GI:3449393  
 VERSION A1100683.1  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsids.  
 1 (bases 1 to 364)  
 Newman,T., deBruijn,F.J., Green,P., Keesstra,K., Kende,H.,  
 McIntosh,L., Ohlrogge,J., Ralkehl,N., Somerville,S., Thomasow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 On May 5, 1995 this sequence version replaced g1:797711.  
 CONTACT: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313tcn@lhm.cl.msu.edu  
 The sequence entry for this EST has been reverse complemented and  
 is being submitted in the sense orientation.  
 Seq primer: M13-21.  
 Location/Qualifiers  
 1..364  
 /organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="187H5XP"  
 /clone\_1ib="Lambda-PRL2"  
 /note="Vector: lambda Zip-Lox; site.1: Sal; site.2: Not;  
 lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7

day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRU's lambda Zip-Lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dr primed cDNA. "

BASE COUNT 113 a 80 c 70 g 99 t 2 others

ORIGIN

Query Match 27.7%; Score 306.2; DB 42; Length 364;  
 Best Local Similarity 95.3%; Pred. No. 1.6e-61;  
 Matches 347; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

544 cccacttcttgaggaaattgaatcctctgtcttttaagcgcttaagcaacggattg 603  
 1 CCCATTTCTTTGGGAAATTAAGATCTMT-GTCTTTTACGGCTTAATGACACCGATTG 59

604 accggtccaatcc-tagaagcaatccagcaatcccaagccttaagttgtgacgtc 662  
 60 ACGGGCCCAATCCNTAGAGCACTCAGTCCATCCCAAGCC-TTAAATGTGTATGTCTC 118

663 aagcaatgattgtgtggaacaatcccaacaagaccccttgcacatcccttaaca 722  
 119 AAGCAATGATTGTGTGGAACAATCCCAACAAGACCTTTCTCACAATTCCTTTACA 178

723 gaacttgaagaacaacccggaatgtgagagacggaattactcgtcttgcaagctcga 782  
 179 GAACTTTGAGAACACCGAGGTGTGAGGACCGGAATTAATCGCTTGCAACTACGA 238

783 caactatgcacctgaacaactggcaaaacctgaagaattgggggtgacctt 842  
 239 CACTAAGTGCACCTGAAAAAATTGGCAAAACCTGAAATGAAGATTGGGGGTGACCTT 298

843 gtaagaacattccaccattatcaaatatcacatctatgtgaataagatatatg 902  
 299 GTAAGACACTTCACCACTTATCAAAATATCAACATCTACTATGTAATATATATATG 358

903 tagt 906  
 359 TAGT 362

RESULT 12  
 Z34606 337 bp mRNA EST 11-JUL-1994  
 LOCUS Z34606  
 DEFINITION A1133366 Versailles-VB Arabidopsis thaliana cDNA clone VBVDH02,  
 mRNA sequence.  
 ACCESSION Z34606  
 VERSION Z34606.1 GI:507009  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsids.  
 1 (bases 1 to 337)  
 CNRS.  
 The Arabidopsis thaliana transcribed genome: the GDR cDNA program  
 Unpublished (1996)  
 CONTACT: Desprez T., Amselem J., Chiapello H., Rouze P., Caboche  
 M., Hofte H.  
 INRA Versailles  
 Laboratoire de Biologie Cellulaire  
 Route de Saint-Cyr, 78026 Versailles Cedex, France  
 Email: therry@versailles.inra.fr.  
 Location/Qualifiers  
 1..337  
 /organism="Arabidopsis thaliana"  
 /strain="ecotype Columbia"  
 /db\_xref="taxon:3702"



BASE COUNT	ORIGIN
73	a 98 c 60 g 105 t 1 others

Query Match	27.68;	Score 305;	DB 20;	Length 337;
Best Local Similarity	97.68;	Pred. No. 3e-61;		
Matches 330;	Conservative	0;	Mismatches	6; Indels 2; Gaps 2;

**Dy** 55 ttgctttccctccttgttcgaaatttccttacctcaattcctttcgatt 114  
**Db** 1 tttgccttcctcttggtcagaaaatttccttacctcaattcctttcgatt 60			

OY 115 ccctctcttaaacctccgaagctcacatyggtctcgaaactatcgttggagctcttc 174  
|||||  
|||  
Db 61 ccccctctttaaacctccgaagctcacatggtctcgaaactatgcgggtggagctcttc 120

Oy    175 gcagcttcgtaaacctaacttagcattgaatcaacctggtcgaagaacaaactccgaaggaa    234  
| | | | |  
Db    121 GCACGTTCTGTAAACCCTAACCTTAGCTTGATTCACCGTGGCAGCAACAACCCGA-GA    179

Qy 235 gatgctcttaagctcttcgcgagatttgacagatccaaaccatgtctctccagagctgg 234  
|||  
Db 180 ATGCMNCTCAACGCTCTTCGCCGGAGTTTGACAGATCCAAACCATTCCTCCAGAGCTGG 239

Oy 295 gatccactctgttcaatccctgtacctggttcattgcaacctgttaaccaagaacaacgcg 356  
 |||||  
 Db 240 GATCCACTCTGTGTATCCCTGTACCGGTTCCATGTCACCTGTAAACCAACAACCGC 299

```

0y      355 gtcactcgtc-ggatttggaaatccaacctctg 391
          |||||
Db      300 gTcAcTcGtGgGgATTTGGGAATTCAAACCTCTCTG 337

```

RESULT 13  
A1100682

DEFINITION	33803 Lambda-PRL2 Arabidopsis thaliana cDNA clone 185LXP 3', sequence.
ACCESSION	AF100682

ABSTRACT	11
INTRODUCTION	11
KEYWORDS	EST.
SOURCE	thale cress.
ORGANISM	<i>Arabidopsis thaliana</i>

eudicotyledons; eudicots: Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS  
1 (bases 1 to 346)  
Newman, T., deBruin, F. J., Green, P., Keegstra, K., Kende, H.,  
McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomasnow  
Reisel, E., and Somerville, C.

TITLE  
JOURNAL  
MEDLINE  
95148729  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)

COMMENT: On May 5, 1995 this sequence version replaced g1:79/667.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University

MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168

Email: [22313tcn@bm.ci.msu.edu](mailto:22313tcn@bm.ci.msu.edu)  
The sequence entry for this EST has been reverse complimented  
is being submitted in the sense orientation.  
Seq primer: M13-21

FEATURES	Location/Qualifiers
----------	---------------------

source 1. .348

```
/strain="var columbiana"  
/db_xref="taxon:3702"  
/clone="185L1XP"
```

/note="Vector: lambda Zip-lox; Site\_1: Sal; Site\_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7

grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and

inserts were directionally cloned with Sal-I Not arms using					
oligo dT primed cDNA,	"				
110 a	74 c	66 g	93 t	5 others	
COUNT					

Very Match 26.48; Score 292; DB 42; Length 348;

574 gttttttacgcgttaatgcaacgcgattgacgcgttccaatcc-tagaacactcacgcgca 632  
 matches 320; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

633 atcccaagccttlaagtttttgaqctcgaagcaatbattttgttqgaacatcccaac 692  
14 gttctttttagggcttatgacaaccgattgacggggccaattccntagagcactactgca 73

693 aagcgaacccttgcacatccttacaqgacttgaacacccggaattggaag 752  
74 ATCCCAAGCC-TTAAgTGTgATGATGCTCAAGCAATgATTGTGTGGAGACATCCCAAC 132

133 AAGCGACCTTTTGCTCACATTCCTTTACAGACTTTGAGAACACCCGAGCTTGGAGGG 192

813 cctgaataatcgaagaatttggggggtgacctctatagaagacacttcaccacttatcaaatat 872  
193 ACCGGATTACTCGNTCTTGCAGACTACGACACTAATCAGCACCIGAAAAAATTTGGCAAAA 252

873 cacatctatattgtaataaagatataatattgact 905  
253 ccigaaaaatgacagattgggggggtgacctgtatagaaaccttcacacactttatcaaatat 312

313 CACATCTACIATGTATAAGTATATATATGTAGT 346

LT	14	96	336 bp	mRNA	EST	30-DEC-1997
STATION	15425	1ambda-DNR2	Arabidopsis	thaliana	cdna	17871177
S	H37296					mdm

```
sequence,
SSION H37296
ION H37296.1 GI:906795
BOM
BOMC
```

CE thale cress.  
GANTSM Arabidopsis thaliana  
Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;

eu dicots; Rosidae II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 336)

McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.

JOURNAL Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 On May 8, 1995 this sequence version replaced g1:801143.  
 ENDLINE















GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:15:32 ; Search time 12463.9 Seconds

(without alignments)  
-76.566 Million cell updates/sec

Title: US-09-180-798-24

Perfect score: 981  
1 agtgcgagtaattcagttcg.....tcgcctaattactcaact 981

Scoring table: IDENTITY\_MWC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_da3:\*  
4: gb\_da4:\*  
5: gb\_da5:\*  
6: gb\_da6:\*  
7: gb\_da7:\*  
8: gb\_da8:\*  
9: gb\_da9:\*  
10: gb\_da10:\*  
11: gb\_da11:\*  
12: gb\_da12:\*  
13: gb\_da13:\*  
14: gb\_da14:\*  
15: gb\_da15:\*  
16: gb\_da16:\*  
17: gb\_da17:\*  
18: gb\_da18:\*  
19: gb\_da19:\*  
20: gb\_da20:\*  
21: gb\_da21:\*  
22: gb\_da22:\*  
23: gb\_da23:\*  
24: gb\_da24:\*  
25: gb\_da25:\*  
26: gb\_da26:\*  
27: gb\_da27:\*  
28: gb\_da28:\*  
29: gb\_da29:\*  
30: gb\_da30:\*  
31: gb\_da31:\*  
32: gb\_da32:\*  
33: gb\_da33:\*  
34: gb\_da34:\*  
35: gb\_da35:\*  
36: gb\_da36:\*  
37: gb\_da37:\*  
38: gb\_da38:\*  
39: gb\_da39:\*  
40: gb\_da40:\*  
41: gb\_da41:\*  
42: gb\_da42:\*  
43: gb\_da43:\*  
44: gb\_da44:\*

45: gb\_da45:\*  
46: gb\_da46:\*  
47: gb\_da47:\*  
48: gb\_da48:\*  
49: gb\_da49:\*  
50: gb\_da50:\*  
51: gb\_da51:\*  
52: gb\_da52:\*  
53: gb\_da53:\*  
54: gb\_da54:\*  
55: gb\_da55:\*  
56: gb\_da56:\*  
57: gb\_da57:\*  
58: gb\_da58:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	981	100.0	981	5 A67819	A67819 Sequence 24
2	924	94.2	1106	5 A67817	A67817 Sequence 22
3	841.8	85.8	1063	5 A67825	A67825 Sequence 30
4	742	75.6	789	5 A67821	A67821 Sequence 26
5	356.2	36.3	894	5 A67823	A67823 Sequence 28
6	210.2	21.4	2089	5 A67827	A67827 Sequence 32
7	156.6	16.0	936	8 SB062279	SB062279 Sorghum bic
8	101.8	10.4	1755	8 DC093048	DC093048 Daucus caro
9	101.8	10.4	1815	5 A67797	A67797 Sequence 2
10	96.6	9.8	4604	7 LERPERGE	X95269 L.esculentu
11	83.2	8.5	101284	7 ATAC009991	ATAC009991 Arabidops
12	81.4	8.3	81513	50 AC021198	AC021198 Arabidops
13	80.8	8.2	134402	8 ATAC011765	ATAC011765 Arabidops
14	80	8.1	112529	44 AC015446	AC015446 Arabidops
15	79.8	8.1	100887	50 ATAC011620	ATAC011620 Arabidops
16	79.2	8.0	116944	8 AC008017	AC008017 Arabidops
17	78.2	7.9	108355	50 AC005957	AC005957 Arabidops
18	77.2	7.8	84440	7 AB010075	AB010075 Arabidops
19	76.2	7.7	118086	7 AT66H11	AT66H11 Arabidops
20	75.8	7.7	142418	7 AP000815	AP000815 Oryza sat
21	75.4	7.7	133840	8 AC003981	AC003981 Complete
22	75	7.6	5940	8 OL072727	OL072727 Oryza longi
23	74	7.5	6695	5 A67796	A67796 Sequence 1
24	73.8	7.5	2786	7 AB029327	AB029327 Nicotiana
25	73.8	7.5	3176	5 E12705	E12705 Arabidops
26	73.8	7.5	3176	5 E12705	E12705 Arabidops
27	73.8	7.5	3176	8 ATU47029	ATU47029 Arabidops
28	73.2	7.5	23143	8 F15P11	F15P11 Arabidops
29	72.4	7.4	4081	5 A67815	A67815 Sequence 20
30	72.4	7.4	98471	7 AC012654	AC012654 Arabidops
31	72	7.3	87434	7 AB010698	AB010698 Arabidops
32	71.8	7.3	9424	8 OSU72724	OSU72724 Oryza sativ
33	71.2	7.3	101154	8 ATRP12	AT13295 Arabidops
34	70.2	7.2	116205	50 AC006931	AC006931 Arabidops
35	69.8	7.1	61712	7 AB017061	AB017061 Arabidops
36	68.6	7.0	13441	8 OLRRKLP1	OLRRKLP1 Arabidops
37	68.6	7.0	52717	7 AB019227	AB019227 Arabidops
38	68.4	7.0	2278	8 SBRUK1	Y14600 Sorghum bic
39	68.4	7.0	3573	5 A57133	A57133 Sequence 4
40	68.4	7.0	4000	8 AF053996	AF053996 Lycopersi
41	68.4	7.0	5368	8 U42445	U42445 Lycopersico
42	68.4	7.0	6471	5 A57130	A57130 Sequence 1
43	68.4	7.0	6471	8 U42444	U42444 Lycopersico
44	67.6	6.9	8416	8 OLRRKLP2	OLRRKLP2 Arabidops
45	67.6	6.9	111945	8 ATRFIC12	AT022224 Arabidops

## ALIGNMENTS



RESULT	1				
LOCUS	A67819				
DEFINITION	A67819	981 bp	DNA		PAT
ACCESSION	Sequence 24	from Patent	WO9743427.		
VERSION	A67819				
KEYWORDS	A67819.1	GI:4756641			
SOURCE	unidentified.				
ORGANISM	unclassified				
REFERENCE	unclassified.				
AUTHORS	1 (phases 1 to 981)				
TITLE	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.				
JOURNAL	PRODUCTION OF APOMITIC SEED				
	Patent: WO 9743427-A 20-NOV-1997;				
FEATURES	CIBA GEIGY AG (CH)				
	Location/Qualifiers				

CDS	104.760
-----	---------

ALTAISPLKVVDDSSNDLCGPIPTNGPFAHIPPLQNEENPRLEGPPELLGLASYDNCNT

BASE COUNT	286 a	236 c	180 g	279 t
ORIGIN				

```

Query Match      100.0%; Score 981; DB 5; Length 981;
Best Local Similarity 100.0%; Pred. No. 1.5e-231;
Matches 981; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY	1	agttggaataatgaattgcttcctccctcttgctcgaagaatttcccttcactca	60
OY	1	agttggaataatgaattgcttcctccctcttgctcgaagaatttcccttcactca	60
Db	1	AGTGTAGATTAATTAGTTGCTCTTCCCTCTTGTTCGAAAATTTCCTTTACTCTCA	60
OY	61	aattccttttcgatattccctctcttaaacctccgaagctcacatgycgtctcgaacta	120
Db	61	AATTCCTTTTGATTTCCCTCTTAAACCTCCGAAAGCTACATAGGGGTCTCGAAACTA	120
OY	121	tcggctgggaagctcttcgcagcttcgfttaacctaaacctagattgattcaacctgfcga	180
Db	121	TGGGTGGAGCTCTTCCAGCTTCGTAAACCTTAACCTTAGCTTTGATTCACTGGTGCGA	180
OY	181	agcaaacctcgaaggaatgctctctcctacgctccttcgccggagttgacagatccagaca	240
Db	181	AGCAAACTCCCAAGAGAAATGCTCTCTACACCTCTTCCGCGAGTTTGACAGATCCAGACA	240
OY	241	tgctctccagagctgggaattccaaactcttgtaacctgttacctggtgtccatgfcacagt	300
Db	241	TGCTCTCAGAGCTGGGATTCAAACTCTGTAACTCTGTACTCTGTTGATTCATGTCACCTG	300
OY	301	taaccagaagaacacgcgtcactcgtgtggaatttgggaaattcaaacctctcttgacact	360
Db	301	TAACCAAGACACACGCGCTCACTGCTGGGATTGGGAAATTCAAACTCTCTGGACATCT	360
OY	361	tgccgcttgagcttgggaagcttgaacattacagttactaagagctctcaaaaaacaacat	420
Db	361	TGCGCTTGAGGTTGGGAAGCTTGAACATTACAGTTCTAGAGCTCTCAAAAAACAACAT	420
OY	421	ccaaggaactatlaacctcgaacttggaaatcgaagaatccatcacagcttgaatctgta	480
Db	421	CCAAGAACTATTAACCTCCGAACTGGGAAATCGAAGATCTCATCAACTTGGATCTGTGA	480
OY	481	caacaacaacttcaaggaatagttccaccattcttgggaaattgaagctctgctctc	540
Db	481	CAACAACAATCTTACAGGAGTAGTCTCCAACTTTTGGGAAATTTGAAGTCTCTGAGCTTT	540

QY	541	tttaagccttaatgcaacccgattgacccgttccaatccctagaacatcagcaatcc	600
Db	541	TTTACGGCTTATGACAAACGATTTGACGGTCCAATCCTTAGAGACTACAGCAATCC	600
QY	601	aagccttaaaagtltgtgacgctccaaagaatgtttgttggaaacatcccaacaag	660
Db	601	AAACCTTAAAGTGTGTGACGCTCCAAACCAATGATTTGTGTGAAACAAATCCAAACACGG	660
QY	661	acccttgcatacttccctttacagaactttggaaacacccgagatcttgagggagccg	720
Db	661	ACCCCTTGCTACATTCCTTTACAAACCTTTGAAACAAACCGAATTTGAGAGACCGGA	720
QY	721	attactcggctcttgcaagcctcagacatactgaccttgaacaacatgycgaacactgaa	780
Db	721	ATTACTCGGCTCTTGCAACACTACGACTAATCTGACCTGAAACAACTGCAAAACCTGAA	780
QY	781	aatgaagaattcggggggtgacccctgttaagaacacttcaacccttatacaataccaatc	840
Db	781	AAATGAAGAATTGGGGGGGACCTCTGTAAACACTTCACACTTTATCAAAATATCACATC	840
QY	841	tattatgtaataagfatatactatgttagtaaaacaacaaaataagaagatcgaaatcgta	900
Db	841	TATTATGTAATAAGTATATATATGTAGTAAACCAAAAAAATGAAGATCGAATCGGTA	900
QY	901	atatcatctgcttcctaatttggaaacttcgagctcgtatgtaanaatttctaattcgatt	960
Db	901	ATATCATCTGCTTCCTAATTTGAGAACTTCGAGCTGTATGTAATAATTTCTAAATGCGATT	960
QY	961	ttgcgcttaataattactccaact 981	
Db	961	TTGCCTTAAATTAATCTCACACT 981	

RESULT	2				
LOCUS	A67817				
DEFINITION	A67817	1106 bp	DNA		
ACCESSION	Sequence 22	from Patent	WO9743427.		
VERSION	A67817				
KEYWORDS	A67817.1	GI:4756639			
SOURCE					
ORGANISM	unidentified.				
	unidentified				
	unclassified.				
REFERENCE	1 (bases 1 to 1106)				
AUTHORS	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.				
TITLE	PRODUCTION OF APOMORPHIC SEED				
JOURNAL	Patent: WO 9743427-A 20-NOV-1997;				
	CIBA GEIGY AG (CH)				

```

/organism="unidentified"
/db_xref="taxon:32644"
142..798
CDS

```

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42244.1"
/bi_xref="GI:4756640"
/translation="MASRWKELFASLTLTALIHVEANSEGDALVLAIRSLTDP
DHYLSMPTLAPNCPCTWENHCNODNRYTVVDGNSNLSGLAPELGRIEHOYLEYEL
KNNGISTIPSELGNKRLISDLTNNNELTIYVTFGRKSKYPRILINDNR TGSILTE
HSRQSAFHYVDVSSNDLCGTLIPNGFAHIPPLQNEENNPRULEGPELLGLASDYINCL
"
BASE COUNT      331 a      258 c      206 g      311 t
ORIGIN

```

Query Match	94.28%	Score 924;	DB 5;	Length 1106;
Best Local Similarity	98.88%	Pred. No. 1.6e-217;		
Matches 952;	Conservative	0;	Mismatches 10;	Indels 2;
Gaps	2;			

  

QY	13	ttagttgccttcctcctcttgcagaaatttccttactctcaaatccctttg	72
I	1		



```

Db 51 TAAATTCGTTTCCTTCCTTTCCTTCAGAAAATTTCCCTTACTCTCAAAATTCCTTTTCG 110
OY 73 attccctctcttaaacctccgaaagctcacatgycgctcgaacatctcgttggagct 132
Db 111 ATTTCCTCTCTTAACCTCCGAAAGCTCACATGGCGCTGAAACATATCGTGGAGCT 170
OY 133 ctctgcaactctgttaaaccttaactgcttattcactgctgctgaagaacatccga 192
Db 171 CTTCGCACTCTGTTAACCTTAACCTTAAGCTTATTCACCTGGTCAAGCAACCTCGCA 230
OY 193 aagagatgctctctacgctctcgcgcgagttgacagatccaaacatgctccacag 252
Db 231 AGGAATGCTCTCTACGCTCTTCGCCGAGTTTACAGATCCACATGCTCCCGAG 290
OY 253 ctggagatccaaactctgttaactcctgtacgtgtccatgctcgttaacgaagca 312
Db 291 CTGGGATCCAACTCTGTTAACTCTTACCTGTTCCATCTCACTGTAACCAAGCA 350
OY 313 cggagtaactgctgtggaattgggaatccaacactcctcgagatcttgcgctgagct 372
Db 351 CCGGCTCACTGCTGTGGATTGGGAATTCAACTCTCTGACATCTTGGCTGAGCT 410
OY 373 tgggaagcttgacaattacagatctagagctctacaacaaacacatccaaggaact 432
Db 411 TGGGAAGCTTGAACTTAACATCTAGAGCTCTACAAAACAACATCCAGGAATAT 470
OY 433 accctccgaacttggaatctgaagaatctatagcttgaattgttaacaaacatct 492
Db 471 ACCCTCCGAATCTGAAATCTGAAGATCTATCAGCTTGATCTGTAACAACAATCT 530
OY 493 taacaggaagtagtcccaactctcttgggaatctgaagctcgttcttcttaagctta 552
Db 531 TACAGGAGATGTTCCACACTTCTTGGGAAATGGAAGCTCTGGCTTTTACGGCTTA 590
OY 553 tgacaaacgaattgaccggtccaaatccctagaagcactcgcgcaatcccaagcc- ttaag 611
Db 591 TGAACAACCAATGACCGGCTCAAT- CCTAGACACTCAAGGCAATCCCAACCTTTAAG 649
OY 612 ttgttagcgtctcaagaactgaattgtgtggaacatcccaacaaacgaaccccttctc 671
Db 650 TTGTTGACGCTCAAGCATATGTTGTGGGACATCCCAACAAGGACCTTTCTCTC 709
OY 672 acatccctcttaagaacttgagaacaaacgagatttggaggagcgaattactcgtc 731
Db 710 AACTTCCTTTACAGAACTTTGAGAACCCGAGATTGAGGAGACCGGAATTAATCTCGGTC 769
OY 732 ttgcaagctagcagactaactgacactgaacaaactgcaaaccttgaaatgaagaatt 791
Db 770 TTGCAAGCTACGACACTGACACTGAAACACTGCAAAACCTGAAATGAAAGATT 829
OY 792 ggggggtgacctgttaagaacacttcaacactttaacaatacacatctatgtlaa 851
Db 830 GGGGGGTGACCTTGTAGAACACTTCAACCTTATCAAAATATCATCTTATATGTAAT 889
OY 852 aagtatataatgtatgaataaacaacaaatgaagaatgtaattgtaatatatctg 911
Db 890 AAGTATATATATGATTAATAAACAATAAATAAAGAAATCGAATGGTAATATCATCTGG 949
OY 912 tctcaattgagaactcgaagctcgtatgtatgaataatttcaaatgcaatttcgccta 971
Db 950 TCTCAATGAGAACTTCGAGGCTCTGTATGTAATAATTTCAATGAGATTTTCGGTACTG 1009
OY 972 tact 975
Db 1010 TAAAT 1013

```

```

RESULT 3
LOCUS A67825 1063 bp DNA
DEFINITION Sequence 30 from Patent WO9743427.
ACCESSION A67825
VERSION A67825.1 GI:4756647

```

```

KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1063)
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
TITLE PRODUCTION OF APOMICRIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
FEATURES
location/Qualifiers
source 1..1063
/organism="unidentified"
/db_xref="taxon:32644"
106..762
CDS
/note="unamed protein product"
/codon_start=1
/protein_id="CAB42252.1"
/db_xref="GI:4756648"
/translation="MASRNYRMELFASLIILALILHLVYANSSEGDALVALRS/ITDP
DHYLOSMDPTLVNRCWPHVYTONODRVRVRLDLSNLSNGHAPLGLKLEHLOYLEY
KNNIQGTIPSELGNLKNLISLDLYNNNLGTYPTSLGKLSLYPLRLDNRNLGPIPR
ALTRIPSLKVVDSNDLGGTITPTNPFAPHLPLQNFENPRLEGPGLGLASTYDICT
BASE COUNT 313 a 242 c 206 g 302 t
ORIGIN
Query Match 85.8%; Score 841.8; DB 5; Length 1063;
Best Local Similarity 96.1%; Pred. No. 2,8e-197;
Matches 885; Conservative 0; Mismatches 32; Indels 4; Gaps 2;
OY 57 ctcaattcccttccatctccctctcttaacctccgaagctccatgctgcga 116
Db 59 CTCAATTCATATGATTAATCTCTCTGACCTCCGATAGCTCAATGCGCTCGAA 118
OY 117 actatcgttggagctcttcgacgttcgttaaccttaaccttaagcttgattcacttg 176
Db 119 ACTATCGGTGGAGGCTCTCCAGCTTCGCACTTCGTTAACTCAACCTTACCTTACCTCG 178
OY 177 tgaagcaactccgaagagatgctctctacgctcttcgacgagttgacagatccag 236
Db 179 TCGAAGCAAACTCCGAAGAGATGCTCTTACGCTCTCCCGGAGTTTAAAGATCCGG 238
OY 237 accatgctcccgagagctcggatccaaactctgttaactctgtacctgttccatgtca 296
Db 239 ACCATGTTCTCCAGAGCTGGGATCCAACTCTTATCTTATCTTACCTGTTCCAGTCA 298
OY 297 cctgttaacaaagacacacgctcactcgtgtggtttgggaattcaaacctctcggac 356
Db 299 CCTGTAAACCAACACCGCGTCACTGCTGTGGATTGGGAATTCAAACTCTGTGGAC 358
OY 357 atcttgcgctgagcttgggaagcttgacaacttaccaglaactagagctctacaacaa 416
Db 359 ACTTGGCGCTGAGCTTGGGAAGCTTGAACTTACAGTATCTTAGACCTGTACAAAAACA 418
OY 417 acatcaaggaactatcctccgaacttggaaactcgaagaactcacaactcagcttgatc 476
Db 419 AATCCAAAGAACTATACCTTCCGAACCTTGAAATCTGAAGAACTCATAGCTTGGATC 478
OY 477 tgtacaacacacatcttcaaggaatgttccactctcttgggaaatltgaatctcgtg 536
Db 479 TGTACAAACAATCTTACAGGAGATGTTCCACTTCTTTGGGAAATTAAGTCTCTGG 538
OY 537 tcttttacggttaatgaacacgattgacgggtccaatccatagagaactaaggcaa 596
Db 539 TCTTTTACGGCTTATATGACAAACGATGAGGGGCCAATCCCTAGGCACTCACTGCAA 598
OY 597 tcccaagcctlaaagtgttgaacgtcgaagcaatgttttgggaacatccacaacaa 656
Db 599 TCCCAAGCCTTAAATTTGTATGTCTCAACCAATGATTTGTGGACAATCCCAACAA 658
OY 657 acggacccttctcacatctccttaacgaacttggagaacacccgagattggaggac 716

```



Db	659	ACGACACCTTTTGCTACACATTCCTTTACAGAACTTTGAGAACACCCGAGGTGGAGGAC	718
Qy	717	CGGAATCtactcgytcccttcgcaagctcaagacactactctgcaccttgaaacaacttgcgcaaac	776
Db	719	CGGAATTCCTCGGCTCTTGCCAAAGCTACACACTAACTCTCCCTGAAAAAATTTGGCAAAACC	778
Qy	777	tgaaaatgaagaatctggggggtgagcccttgtaaagacacttcacacactttcaaatata	836
Db	779	TGAAATGAAAGAAATTTGGGGGGTGACCTTTGTAGAAACACTTCACCACTTTATCAAAATATCA	838
Qy	837	cactctatctgtaataagtatatatatagtgtgataaaccaaaaaaataatgaagaatctgaatc	896
Db	839	CATCTACTATGTATTAATAGTATATATATGTATGTCCTCA--AAAAAAATGAAGAAATCGAATC	896
Qy	897	ggttaatatcactctggtctccaattagaacctcgagtc--tgtatgtanaaatcttctaact	954
Db	897	AGTAATATCATCTGGTCTCATATTCAGAACTTTGGAGCTGTGTATGTATAAATTTCTTAAT	956
Qy	955	ggcgtatttcgcgctaataact 975	
Db	957	GCGACTTTCGCTACTGTAAT 977	

[illegible]

CDS	BASE COUNT	ORIGIN
<pre> /organism="unidentified" /db_xref="taxon:32644" &lt;2..664 /note="unnamed protein product" /codon_start=1 /protein_id="CAB42248.1" /db_xref="gi:4756644" /translation="RPTPRARYRWELPAASLIITLAIHIVEANSEGDALVALRRSIT DPDHLQSMDEPTLVNCPCTWAFVTCENQDRVRYVDLGSNSLGHAPLEGKEHLYQYEL LYKNNGIGTIPSELGNLKNLISDLYNNNLGIPTSGIKRKSIVFRLINDNRITGPT PRAITAFPSLKVVDVSSNDLCGITPTNGPFAHIFLQNFENNPRLEGELGLASVDIN CT" </pre>	<p>235 a 191 c 156 g 207 t</p>	

Query Match	75.68;	Score 742;	DB 5;	Length 789;
Best Local Similarity	96.88;	Pred. No. 1.1e-172;		
Matches 757; Conservative	0;	Mismatches 25;	Indels 0;	Gaps 0;

Qy	101	cacatagcgctcgcgaacatcgtgagggccttcgcagcttcgttaacctaacctta	160
Db	5	ccacagccgtccgcgaacatcgtgagggagccttcgcagcttcgttaacctaacctta	64
Qy	161	gcttggattcaacctggtcgaagaacatccgaagagatgctctctaagctcttcgccgg	220
Db	65	ggcttgattcactctgctccaagcacaactccgaaggagagctttacacctttgcgccgg	124
Qy	221	agtttcacagatccacagacatgtctccagaagctggatccaacctgttaactccgt	280
Db	125	agtttaacagatccacagacatgtctccagaagctggagccaactcttttaactccttct	184
Qy	281	acctggttcacatgctcaacctgttaaccaagaacacgcgtcaactcgtgttgattggaaat	340

Db	185	ACCTGGTTCACGTACCTGTAAACCAAGAACACGGGTCACTCGTGTGATTTGGGAAAT	244
Oy	341	tcgaacctctc tggacatcttggccctggaccttgggaaccttgaacattacagta	400
Db	245	TCAAACTCTTGAGCAATCTGGCCCTGGACCTTGGGAACCTTGAACATTTCACTATCTA	304
Oy	401	gaacctacaanaaacaatccaaagacacataacctccgaactctggaatctgaagat	460
Db	305	GAGCTTACAAAACACATCCAGGAACATACCTCCGAATCTGGAAATCTGAAGAT	364
Oy	461	ctatacagcttggatctgtatacaacaacatcttaacaggaatagtctccactcttggga	520
Db	365	CTATACAGTTGGATCTGTACACAAACATCTTACAGGATGATGTTCCACATCTTTGGGA	424
Oy	521	aaattgaagctctgtctcttctttagcgcttaatgaaacacgaattgacccgtccaatccct	580
Db	425	AAATTGAAGTCTCTGCTCTTTTACGGCTTAATGACACCGATTAACGGGGCCAAATCCT	484
Oy	581	agagcacctcaacggcgaatcccaagccttaaaagtgttgaagcttcaagaatgatttgt	640
Db	485	AGAGCACTCACATCGCATCCCAAGCCTTAAAGTTGTTGATGTCTCAAGCAATGATTGTGT	544
Oy	641	ggaacatccccaacaaagggaccccttggccacatcccttacaagacttggagaaac	700
Db	545	GGAACATCCCAACAAAGGACCTTTTGTCACATTCCTTTACAGACTTTGAGACAC	604
Oy	701	ccgaagattggaaggaccggaattactcgtcttggcaagctacgaacactaactgcactga	760
Db	605	CCGAGGTTGGAGGAGCCGGAATTACTGGTCTTGGCAAGTACGACACTTAACATGCACTGA	664
Oy	761	aacacactggaaaacctaagaatgggaatgggggtgtaaccttgaagaacactcaac	820
Db	665	AAAAATTGGCAAACTGAAAAATGAAGATTGGGGGTGACTCTTGAAGAACACTTCACC	724
Oy	821	acttatcaaatatcacatctatattgtaabaagtatatatatgtagtgaataaacaanaa	880
Db	725	ACTTATCAAAATTCACATCTACTATGTAATAGTATATATATGTAGTCCCAAAAAAAA	784
Oy	881	aa 882	
Db	785	AA 786	

```

RESULT      5
A67823
LOCUS       A67823               894 bp    DNA
DEFINITION  Sequence 28 from Patent WO9743427.
ACCESSION   A67823
VERSION     A67823.1  GI:4756645
KEYWORDS
SOURCE      .
ORGANISM    unidentified.
            unclassified.
REFERENCE   1  (bases 1 to 894)
            De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.
            PRODUCTION OF APOMITIC SEED
            Patent: WO 9743427-A 20-NOV-1997;
            CIBA GEIGY AG (CH)

FEATURES
    source   Location/Qualifiers
             1..894
             /organism="unidentified"
             /db_xref="taxon:32644"
             <1..678
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="CAB42250.1"
             /db_xref="GI:4756646"
             /translation="GPIQASEGDALHMLRSLSDPDNVYVSWDPFLVNPCTWEHVTCN
             OHQWTRLDGNSMLSGHVLPELGLHLYGIIILPFDYIKTFLSTATHTFCF
             EASYSLVNEIQGTIPSELGNLKLISIDILNNMLTGKISSLGKLSLVLRNENR
             LTGPRLPRLTVISSLKVYDVSGNDLCGTIPVEGFPHIIPQNFENNRLRGPELGLA
             SYDINCT"

```







```
RESULT 7
SB062279 936 bp mRNA PLN 09-DEC-1996
LOCUS Sorghum bicolor leucine-rich repeat-containing extracellular
DEFINITION glycoprotein mRNA, complete cds.
ACCESSION U62279
VERSION U62279.1 GI:1710123
KEYWORDS
SOURCE
ORGANISM sorghum.
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum.
1 (bases 1 to 936)
Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.
Isolation of a cDNA encoding a novel leucine-rich repeat motif from
Sorghum bicolor inoculated with fungi
Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)
JOURNAL 97124217
MEDLINE Erratum: [[published erratum appears in Mol Plant Microbe Interact
REMARK 1997 Mar;10(2):3021]]
2 (bases 1 to 936)
Hipskind,J.D.
REFERENCE Direct Submission
AUTHORS Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant
JOURNAL Pathology, Purdue University, West Lafayette, IN 47907, USA
TITLE Location/Qualifiers
FEATURES
SOURCE 1. 936
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone="SLR8"
12. 755
/function="protein-protein ligand binding"
/note="leucine-rich repeat-containing extracellular
glycoprotein; contains six N-glycosylation sites
[MX(5/7)]"
/codon_start=1
/protein_id="AAC49559.1"
/db_xref="GI:1710124"
/translation="MAPQFAAGFLTGILAFASCNREGILYKRLAWDPNNVL
/OSMSTLANPCTFHVTCNNNNVIRVDGNATISPLPDLAEIONLQYIELYGL
NGSIPETGLNLTNLSLDNDLGEIPTLRYLRKYONNLTGPPSSFGN
LTSLESKLOENSLGASIPASLNIKALQFSRLNDMLTGTVPKSFSTGNTLTEL
NTDRNNLDGRTSSGLRVTALIIDALKTA"
CDS
s1g_peptide 36. 77
/evidence="not experimental"
BASE COUNT 247 a 232 c 207 g 250 t
ORIGIN
Query Match 16.0%; Score 156.6; DB 8; Length 936;
Best Local Similarity 58.5%; Pred. No. 1.6e-28;
Matches 273; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
```

```
QY 485 aacaacttacaaggatagttccactcttcttggaanaattgaagtcctgcttttta 544
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 AACCTTACCAGGCGGAATCCCAATACGCTTGCTCTGACGACGCGCATATCTG 440
QY 545 cgccttaatgcacacgcattgaccgcgcgaatccctagagcatcaacgaatcccaagc 604
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 441 AGGTTGACCAAAACAACTGACAGGCGCTATACATCATCTTTGGCAACTGACTAGC 500
QY 605 cttaaatgttgtagcgtctcaagcaatgttgtyggaacaatccc 651
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 501 CTTGCGAATGAGAGCTTCAAGAGATTCGTTAGCGCGCATATTC 547
RESULT 8
DCU93048 1755 bp mRNA PLN 28-JUN-1997
LOCUS Daucus carota somatic embryogenesis receptor-like kinase mRNA,
DEFINITION complete cds.
ACCESSION U93048
VERSION U93048.1 GI:2224910
KEYWORDS
SOURCE carrot.
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
1 (bases 1 to 1755)
Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
A leucine-rich repeat containing receptor-like kinase marks somatic
plant cells competent to form embryos
Development 124 (10), 2049-2062 (1997)
JOURNAL 97313247
MEDLINE 2 (bases 1 to 1755)
REFERENCE Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
AUTHORS Submitted (12-MAR-1997) Molecular Biology, Agricultural University
JOURNAL of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
TITLE Location/Qualifiers
FEATURES
SOURCE 1. 1755
/organism="Daucus carota"
/db_xref="taxon:4039"
34. 1695
/note="SERK"
/codon_start=1
/product="somatic embryogenesis receptor-like kinase"
/protein_id="AA861708.1"
/db_xref="GI:2224911"
/translation="MNRNSINILNYQFTDAYLDKYGVLMTLELYSNNISGPIPSDLG
NLTNIVSLIDLYMNSFSGPIPDITGKLTREFLRANNSLSGPIPLSLTNITLQVLDL
SNRLSGPVPDNGSFSLETPISFANNILCGPYTGRRPGSPFSPFPPIPTVOP
PGONGPGIAGAGAGALLFPAAALFAFMWRKRRPREFPVEPPEEVLGGLK
RFSRLQVAVTDFPSTILGSGGKGYKAAKFAKFAVYKRLKEETPGSELQFOREV
EMTSMAYHNRNLRLRGFCMPTERLVIYPMAGSVASCSLRQSPSEPLDPTPRRI
ALGSAKGLSTLHDCDPKTIHROVKANILDEEFAVVDGRLALNDYKDTHTYTA
VRGTGLYIAPETVLSGKSSBKTFVFGIMLLELTIGQRAFDLARIANDDDWLLDWV
KSLKEKRLKMLVDPLENNYIDTEVQLIYVALICTOGSPMERPKMSVRYRLBEDG
LAERKDEMOKEVYHDOVELAPRHSYMIIDSTINDINHAFLSGPR"
BASE COUNT 506 a 347 c 407 g 495 t
ORIGIN
Query Match 10.4%; Score 101.8; DB 8; Length 1755;
Best Local Similarity 55.1%; Pred. No. 5.1e-15;
Matches 199; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
```



QY	481	caacaaacatcttaaggaatagctcccaactctcttggaaaaatgagctctcgtgctt	540
Db	198	CATGAATACGCTCTCTGGACCTATACCGGACACATTAGAGAGCTTAAAGCTTAAGATT	257
QY	541	tttagcgcttaatgacaacgcattgacccggttccaatcccttagagactacagcgcaatcc	600
Db	258	CTTGCGCTCAACAACAACAGCGCTCTGGGTCGAATTCAGTACGTACATTAATTATAC	317
QY	601	aagcgcttaagttgttgacgcgtcaagcaatgatttgytggaaaaatcccaaaaacgg	660
Db	318	AACCTCTCAAGTCCTGGATTATCAACAACATCGGCTATACGACCAAGTACCGGATTAATGG	377
QY	661	accctttgctcaatctctcttaacgaacttggaaaacaacccgagattggaagggaccgga	720
Db	378	CTCATTTCTCTGTGTACACCATATCAGTTTGGCAATTAATTGAATTATATGGACCGGT	437
QY	721	a 721	
Db	438	A 438	

QY	481	caacacaaatctctaaaggatagctcccaactctcttggaagaatgatgctctcgtgctt	540
Db	258	CAGGAATACCTTTCTTGAGACCTATACCGGACACATTAGGAAGCTTTAAAGGCTTAAGATT	317
QY	541	tttagcgcttaatgacaacccgattgaccggtgccatccctagaagcactcagcgaaatccc	600
Db	318	CTTGCGCTCAACACACACACACGCTCTGTGGTCAATTCAATGTCAAGTCACTAATATTAC	377
QY	601	aagccctaaagtctgtgcgcgtctaaagcaatgatcttgctgtgaaacaatcccaacaacg	660
Db	378	AACCTTCACAGTCGAGATTATATAAACAATGGCATATAGGACCAAGTACCGGATAATGG	437
QY	661	accctctgtgcacatctctcttaccgaactcttgaaacaacccgagatcttgaaaggaccga	720
Db	438	CTCATTTCTTGTTTAAACCTATACAGTTTGGCATAATATTGAATTTATGTGGACCGGT	497
QY	721	a 721	
Db	498	A 498	

RESULT	9				
LOCUS	A67797				
DEFINITION	A67797	1815 bp	DNA		
ACCESSION	Sequence 2	from Patent	WO9743427.	PAT	05-MAY-1999
VERSION	A67797				
KEYWORDS	A67797.1	GI:4756623			
SOURCE	carrot.				
ORGANISM	Daucus carota				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.				
AUTHORS	De V.S., Schmidt, E.D., Van H.G. and Recht, V.F.				
TITLE	PRODUCTION OF APOMICRITIC SEED				
JOURNAL	Patent: WO 9743427-A 20-NOV-1997;				
FEATURES	CIBA GEIGY AG (CH)				
source	Location/Qualifiers				
	1..1815				

RESULT 10			
LETRGENE	LETRGENE	4604 bp	DNA
LOCUS	L.esculentum LRP gene.		
DEFINITION	X95259		PLN
ACCESSION	X95259.1	GI:1619299	10-OCT-1996
VERSION	LRP gene; LRP protein.		
KEYWORDS	tomato.		
SOURCE	Lycopersicon esculentum		
ORGANISM	Eukaryota; Vitidiplantae; Streptophyta; Embryophyta; Tracheophyta; euPharophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; Solananceae; Solanales; Solanaceae; Solanum; Potatoes; Lycopersicon.		
REFERENCE	1 (bases 1 to 4604)		
AUTHORS	Vera, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-JAN-1996) P. Vera, Universidad Politecnica de Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino de Vera 14, E- 46022 Valencia, SPAIN		

```

/organism="Daucus carota"
/db_xref="taxon:4039"
94. .1755
/note="unnamed protein product"
/coodon_start=1
/coodon_id="CAB4241.1"
/db_xref="GI:4756624"
/translation="MNRNSINLWYMOFTDAYLDYKGYLMTLETYSNNISGPIPSIDG
NITNYSIDLYMNSFSPIDPTGKRLTLEPIPLNNSISGSIPLSINITITVDIL
NNRSLGSPYDNGSPSFLTEPISPANNINLTCGPTGCGSGSPFPEPPPIIPETSTVP
PQONPPTCAINGVYAGAGALLPSPANNNLTPGKRLTLEPIPLNNSISGSIPLSINITITVDIL
KPSRLQVLYDITSTILDRGFGSKVYKGLASGVAKLRERSTGGELQGTVEI
NLSAVERLNLRLRGCPMTERTLVTPYMAKNSILDFEEAAYGSGLARMDYKDTVTA
ALGSAAGSYLHBCDPRILHRDKANILILDEFAVAGDGLARMDYKDTVTA
VGTSGYLAPEYLSGKSEKTDVFGYINILITGQAFPLARLADDVYMLDWW
KSLTLERKLEMLVDEDLNENITDIEVEQLQVALLCTQGSPEPERKMEVYMLDEGGQ
LAEKDEWQKVEYHQDETLAPDSTNMIIDSTNMLAEELGSPR"
531 a 354 c 415 g 515 t

```

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	FEATURES
2 (bases 1 to 4604)	Tornero,P., Mayda,E., Gomez,M.D., Canas,I., Conejero,V. and Vera,P.	Characterization of LRP, a leucine-rich repeat (LRP) protein from tomato plants that is processed during pathogenesis	Plant J	10 (2), 315-330 (1996)	96367673
					Location/Qualifiers
					1..4604
					/organism="Lycopersicon esculentum"
					/cultural="VF8"
					/db_xref="taxon:4081"
					/tissue_type="leaf"
					<1..345
					/number=1
					113..4355
					/gene="LRP"
					join(113..345,2381..2452,2565..2708,2928..2999,3081..3152,4283..4355)
					/gene="LRP"

[illegible]

intron  
 /gene="Lrp"  
 /number=1



```

exon      2381..2452
           /gene="LRP"
           /number=2
Intron    2453..2564
           /gene="LRP"
           /number=2
exon      2565..2708
           /gene="LRP"
           /number=3
Intron    2709..2927
           /gene="LRP"
           /number=3
exon      2928..2999
           /gene="LRP"
           /number=4
Intron    3000..3080
           /gene="LRP"
           /number=4
exon      3081..3152
           /gene="LRP"
           /number=5
Intron    3153..4282
           /gene="LRP"
           /number=5
exon      4283..>4604
           /number=6

BASE COUNT 1311 a 758 c 888 g 1647 t
ORIGIN

```

```

Query Match      9.8%; Score 96.6; DB 7; Length 4604;
Best Local Similarity 69.1%; Pred. No. 9.9e-14;
Matches 132; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```

```

QY 140 gccctgaaccttaaccttagcttgatccaccctgctggaagaacacacccgaagagat 199
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 158 GGTGTGTTGGCAGTGTGCTTCTGCTAGAGCTGTGCTGAAGGGGAATTCAGAGGGGAT 217

QY 200 gccctgaacctcttcgacgagattgacagatcacagacacatctctcgaagcttgagt 259
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 218 GCTTTGACGCCCTTCGCCGACCTATCTGACCCGGGTAACGTGTACAGACTGAGAT 277

QY 260 ccaacctctgtaacctctgtaacctggtccatgtaacctgtaaccaagaacacgcgcgc 319
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 278 CCAATCTTTTAACTTACTGTTCACTGTTCACTGTTCACTGTTCACTGTTCACTGTTCA 337

QY 320 actcgtgtgga 330
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 338 ACTCGTGTGTA 348

```

```

RESULT 11
ATAC009991/c
LOCUS ATAC009991 101284 bp DNA PLN 08-OCT-1999
DEFINITION Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence,
complete sequence.
ACCESSION AC009991
VERSION AC009991.3 GI:6016677
KEYWORDS HTG
SOURCE thale cress
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          euophyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
          eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
          Arabidopsi
REFERENCE 1 (bases 1 to 101284)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B.,
          Rensing,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
          Bowman,C.L., White,O., Niernan,W.C. and Fraser,C.M.
TITLE Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 101284)
AUTHORS Lin,X. and Kaul,S.

```

```

TITLE Direct Submission
JOURNAL Submitted (09-SEP-1999) The Institute for Genomic Research, 9712
REFERENCE Medical Center Dr., Rockville, MD 20850, USA, xlinetlgr.org
AUTHORS 3 (bases 1 to 101284)
TITLE Lin,X.
JOURNAL Direct Submission
COMMENT Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
          Medical Center Dr., Rockville, MD 20850, USA
          On Oct 8, 1999 this sequence version replaced gi:5902413.
          Address all correspondence to:
          Xiaoying Lin
          The Institute for Genomic Research
          9712 Medical Center Dr
          Rockville, MD 20850, USA
          e-mail: xlinetlgr.org
          BAC clone F9f8 is from Arabidopsis chromosome III and is near the
          molecular marker g4547.
          The orientation of the sequence is from SP6 to T7 end of the BAC
          clone.

```

Genes were identified by a combination of three methods: Gene Prediction programs including GRAIL (available by anonymous ftp from [arthur.epm.ornl.gov](http://arthur.epm.ornl.gov)), GeneFinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://www.cbs.dtu.dk/netgene/cbsnetgene.html>), and NetPlantGene (<http://www.cbs.dtu.dk/netgene/cbsnetgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tldb/at.html>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted as tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

## FEATURES

```

source          1..101284
                 /organism="Arabidopsis thaliana"
                 /cultivar="Columbia"
                 /db_xref="taxon:3702"
                 /chromosome="III"
                 /map="g4547"
                 /clone="F9f8"
                 /note="exon predicted by xgrail, quality excellent"
                 124..259
                 /note="exon predicted by xgrail, quality excellent"
                 /complement(347..408)
                 /note="exon predicted by xgrail, quality
                 excellent_shadowexon"
                 355..594
                 /note="exon predicted by xgrail, quality excellent"
                 672..752
                 /note="exon predicted by xgrail, quality excellent"
                 854..973
                 /note="exon predicted by xgrail, quality excellent"
                 /complement(1182..1357)
                 /note="exon predicted by xgrail, quality
                 marginal_shadowexon"
                 complement(1516..1694)
                 /note="exon predicted by xgrail, quality
                 excellent_shadowexon"
                 1546..1666
                 /note="exon predicted by xgrail, quality good"
                 1747..1838
                 /note="exon predicted by xgrail, quality excellent"
                 complement(2090..2250)
                 /note="exon predicted by xgrail, quality good_shadowexon"

```



misc_feature	2114..2246	/note="exon predicted by xgfall, quality excellent"
repeat_region	complement(2975..3024) /rpt_family="(CAA)n"	
mRNA	join(3070..3325,3565..3642,4094..4234,4338..4414, 4506..4655,4800..5013) /gene="P9F8.1"	
gene	3070..>5013 /gene="P9F8.1"	
CDS	/note="similar to unknown protein GB:AAC27167 [Arabidopsis thaliana]" join(3215..3325,3565..3642,4094..4234,4338..4414, 4506..4655,4800..5013) /gene="P9F8.1"	
note="unknown protein"	/codon_start=1 /protein_id="AAFO1505.1"	
repeat_region	/db_xref="GI:6016678"	
misc_feature	/translation="MVGAPQIVLFGSSIVQMGFGHGWGALLSEYAKRADIILNG KIALHMSRLAEVDQVPEPKDAVQSLIVYVGGDSMAHPSGLGPHVLEVDNM QELGLEVDLESTFOKADDKTKVCTDGIHLSAGSKIVAGEIILRVKEDMNPSLHM KSMTEPADSDSEYDLSVADGQFYANSEWTFMEQWD"	
repeat_region	complement(5513..5597) /rpt_family="(GA)n"	
misc_feature	complement(5527..5597) /note="exon predicted by xgfall, quality excellent shadowexon"	
mRNA	join(5588..5795,5897..5929,6013..6235,7264..7387, 7467..7932)	
gene	/gene="P9F8.2"	
CDS	5589..7932 /gene="P9F8.2"	
note="similar to nucleic acid binding protein Alfin-1 [Arabidopsis thaliana]"	join(5696..5795,5897..5929,6013..6235,7264..7387, 7467..7727)	
gene	/gene="P9F8.2"	
note="putative nucleic acid binding protein"	/codon_start=1 /product="putative nucleic acid binding protein"	
repeat_region	/protein_id="AAFO1506.1"	
repeat_region	/db_xref="GI:6016679"	
repeat_region	/translation="MAAAVSSNPRTVEIEIKDYSARRAALLRLTKDVDDFYSQCP EKEHLEKIPHEWSEVNLPAEVEPELPEPLGINFARDQMKRDMVSLVAHSDGW LISVSYFARLNRNKEKRLFLINDLPELVYVGRAMDMPSDSGSKNGTK RSIDGQTKSTPKLMESEYEEEDDEHDITLCSGGGHYNEEFWICDCEKWTYHGI KCKYIRPAKESIKYKCCPCPKCAKKGQ"	
repeat_region	complement(7717..7760)	
repeat_region	/rpt_family="(CAT)n"	
repeat_region	complement(7810..7847)	
repeat_region	/rpt_family="(GAAA)n"	
mRNA	8515..8572 /rpt_family="(TAA)n"	
gene	complement(join(9426..9674,10198..10528,11098..11345, 11491..>11865)) /gene="P9F8.3"	
gene	complement(<9426..>11865) /gene="P9F8.3"	
note="similar to leucanthocyanidin dioxygenase GB:BA20143 [Perilla frutescens]"	complement(join(9426..9674,10198..10528,11098..11345, 11491..11865)) /gene="P9F8.3"	
CDS	/codon_start=1 /product="putative leucanthocyanidin dioxygenase"	
repeat_region	/protein_id="AAFO1507.1"	
repeat_region	/db_xref="GI:6016680"	
repeat_region	/translation="NNNLEIKIESKTCINDQEOEYKIDNMMSQDKNKLEIKKSG LGKRWPIYRQVSLAESNLTSLPDYIKPSPORQGTTLIDHPVADINILIDDS LPSGNDKGRKRISEACREMGROYINHGKRPILMAARETMSFENLVEAKVEYNS PRTEBEGDSKRISEAKREMGROYINHGKRPILMAARETMSFENLVEAKVEYNS VLTGKRIEAGKAILDMNDYITLHFLPLALCKDFNKPSPSPENITBMDDEKEL VGTGILRLPTLSSNIGLRAQQLDEAFGGSDVACLRVNYPCQPELALGISPSHSD GKATILPDDQVYQVRHDTWITVNPRAFIYVIGDOIILSNSKSYSEHRYIV	

[illegible]



ORGANISM	Eukaryotes: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta: Euphyllophytes; Spermatophyta: Magnoliophyta: eudicotyledons; core eudicots: Rosidae: eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE	1 (bases 1 to 81513)
AUTHORS	Liu,S.X., Sakano,H., Yu,G., Lee,J.M., Lenz,C., Pham,P., Toriumi,M., Chin,C., Chion,J., Choi,E., Chung,M., Gonzalez,A., Hwang,B., Liu,A., Vaysberg,M., Altafi,H., Brooks,S., Buehler,E., Cho,O., Conn,L., Conway,A.B., Hansen,N.F., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Eckert,J.R., Federspiel,N.A. and Theologis,A.
TITLE	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 81513)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE	3 (bases 1 to 81513)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	On Feb 11, 2000 this sequence version replaced g1:6693723. The sequence is of BAC F14D7 from Arabidopsis thaliana chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 6954 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide small overlaps (200 bp) between overlapping submitted clones. The 5' end of this sequence overlaps by 200 bp to the 3' end of the sequence of the clone F1504.
FEATURES	Location/Qualifiers 1..81513     /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="1" /clone="F14D7"
BASE COUNT	27779 a 13814 c 13868 g 26052 t
ORIGIN	
Query Match	8.3%; Score 81.4; DB 50; Length 81513;
Best Local Similarity	54.1%; Pred. No. 5,6e-10;
Matches 166; Conservative	0; Mismatches 141; Indels 0; Gaps 0
OY	345 accctcttggaacatcttgcgctcgtgagcttgggaagtctgaatttacagtatctaaagc 404
DB	2466 ATCTACTGTGGGCATTCGCCCGAAGCTAGGCAACATAACATCATGATCGATTGAAGT 2525
OY	405 tcatacaaaaacaacatccaagaagacatcatcccttcgaaattggaaatctgaagaatcca 464
DB	2526 TGAATTAACAACAACATTAACGGGTTCATTCCTCTCCATTAGAATTCACAAGATTTGA 2585
OY	465 tcagcttgagatctgtacacaacaacatcttaagagtagttcccactctcttggaaaaat 524
DB	2586 CTATTTCTTAACTTTACGAGAACAATATTTAACGTGCTCATTCGCCCGAAGTAGGCAACA 2645
OY	525 tgaagtcctcggtctcttttcaggctaagaacaacgatgaccggtgccaatccctagag 584
DB	2646 TGGAAATGCAATGATTGATTTCAGTGTAATTAACAACAACATTACCGGTTCATTCCTTCT 2705
OY	585 cactcacggcaatcccaagccttaaaagtgttgagcgtccaagcaatgatttgttggaa 644
DB	2706 CCTTCGGAANAATCTCAAGAACTGACGTATCTTTATCTTACCATGATTAATTAACTGGTG 2765
OY	645 caatccc 651
DB	2766 TCATTCC 2772

RESULT	13
LOCUS	ATAC011765
DEFINITION	Arabidopsis thaliana chromosome I BAC FIM20 genomic sequence,
ACCESSION	AF011765.
VERSION	AC011765.4
KEYWORDS	HTG.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi s. 1 (bases 1 to 134402) Lin.X., Kaul,S., Town,C.D., Benito,M., Greasy,T.H., Haas,B., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M. Arabidopsis thaliana chromosome I BAC FIM20 genomic sequence Unpublished 2 (bases 1 to 134402) Lin.x. and Kaul.S. Direct Submission Submitted (14-OCT-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, xlinet@ig.r.org 3 (bases 1 to 134402) Lin.X. Direct Submission Submitted (08-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA On Dec 8, 1999 this Sequence version replaced gi:6102640. Address all correspondence to: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA email: xlinet@ig.r.org BAC clone FIM20 is from Arabidopsis chromosome I and is near the molecular marker ml425. The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of three methods: Gene prediction programs including GNAI1 (available by anonymous ftp from arthur.gem.ornl.gov), GeneFinder (Phil Green, University of Washington), GENSCAN (Chris Burge, <a href="http://genome.stanford.edu/~chris/GENSCAN.html">http://genome.stanford.edu/~chris/GENSCAN.html</a> ), and NetPlantGene ( <a href="http://www.cbs.dtu.dk/netplantgene/cbsnetplane.html">http://www.cbs.dtu.dk/netplantgene/cbsnetplane.html</a> ), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR ( <a href="http://www.tigr.org/tdb/at/cat.html">http://www.tigr.org/tdb/at.cat.html</a> ). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are name after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical'. proteins. Genes encoding tRNAs are predicted by tRNAscan-SF (Sean Eddy, <a href="http://genome.wustl.edu/eddy/tRNAscan-SF/">http://genome.wustl.edu/eddy/tRNAscan-SF/</a> ). Simple repeats are identified by RepeatMasker (Arian Smit, <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a> ). Regions of genomic sequence that are not annotated as genes but have predicte exons by GNAI1 are annotated as misc features. Location/Qualifiers 1..134402 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="I" /map="ml425" /clone="FIM20" 138..269







```

Db 12163 GMAAGACGCGGGGCTTCAAGCACTGACCTGATCTCTCTTTAAACGAAGTCAGCGGTTGCATAC 12222
Oy 507 ccactctcttgggaaatgaagtcctcgtctcttcttaacgcttaatgaacaccgatitga 566
Db 12223 CAGGTCATTATGGGAATATGACCTCTCTTTTGTGGCTAATGCTTGCAACAACTCTTAT 12282
Oy 567 ccggtccaatcccttagagcactcaagcgcgaatcccaagcottaagaattgtagctccaa 626
Db 12283 CAGAGAGAAATTCCTCCAGAGATTTGGCAACTGCACAAAGCCTTTTGTGGTTAACTGGCAA 12342
Oy 627 gcaatgattctgtgtggaacaatcc 650
Db 12343 ACAACCAAGCTCTGCTGATGATTC 12386

RESULT 14
AC015446/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosid II; Brassicales; Brassicaceae;
Arabidopsi s
REFERENCE
AUTHORS
1 (bases 1 to 112529)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altati,H., Nguyen,M., Lam,B., Southwick,A., Miranda,M., Brooks,S.,
Buehler,E., Chao,Q., Chin,C., Chlou,J., Choi,E., Gonzalez,A.,
Hwang,B., Johnson-Hopson,C., Khan,S., Kim,C., Koo,T., Lee,J.M.,
Lenz,C., Liu,A., Liu,S., Mukharsy,N., Pham,P., Sakano,H.,
Shim,P., Tortum,M., Vaysberg,M., Yu,G., Ecker,J., Theologis,A.
and Davis,R.W.
Unpublished
2 (bases 1 to 112529)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altati,H., Nguyen,M., Lam,B., Southwick,A., Bel,Q., Buehler,E.,
Chin,C., Chlou,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B.,
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
Mukharsy,N., Pham,P., Sakano,H., Schwartz,J., Shim,P.,
Thavel,I.A., Tortum,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
Theologis,A. and Davis,R.W.
Direct Submission
Submitted (16-NOV-1999) DNA sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT
On Jan 12, 2000 this sequence version replaced gi:6437511.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3149: contig of 3149 bp in length
* 3150 3199: gap of unknown length
* 3200 12898: contig of 9699 bp in length
* 12899 12948: gap of unknown length
* 12949 50756: contig of 37808 bp in length
* 50757 50806: gap of unknown length
* 50807 112528: contig of 61723 bp in length.
*
* Location/Qualifiers
* 1. 112529
* /organism="Arabidopsis thaliana"
* /db_xref="taxon:3702"
* /chromosome="1"
* /clone="F12G12"
FEATURES
source
35310 a 20450 c 20573 g 36044 t 152 others

```

Query Match	8.2%	Score 80;	DB 44;	Length 112529;
Best Local Similarity	53.9%	Pred. No. 1,3e-09;		
Matches 164;	Conservative 0;	Mismatches 140;	Indels 0;	Gaps 0;
Qy 327	tggaattgggaatcaaacctctctgacatctctgcgcctgagcttgggaattggaac	386		
Db 86684	TAGAGTGCACACACACACATCACATCACCGAGATATTTCTCGTCAGTGGGAAACCTTGGA	86625		
Qy 387	attacaglatctagagctctacaanaacaacatccaaggaactataccctcgaaattg	446		
Db 86624	ATCTGAGACACACCTTGATTGTGACAGAAACACCTTCACAGGAACATATACCGCTGAGTTTG	86565		
Qy 447	gaactcgaagaatctctgagctgagctgcgtcaacaacaactcttcaggagatgttc	506		
Db 86564	GGAACTTAAGCTACTTTAAACAGCTGATCTTCACACACATCTCTCTACTGTCATAATCC	86505		
Qy 507	ccactctctgggaatctgagctctgctgctcttcttcaggacttaatgacacagattga	566		
Db 86504	CBAATCTATCAAGAAATCTACAAAAGCTTACACTCTTGATCTTACGACCAAGCTTAT	86445		
Qy 567	ccggtccaatccctcagagcactcaagggaatcccaagccttaagttgttagctctcaa	626		
Db 86444	CCGGTGAATCCACACAGAACTCGGTCAAGTCACAGCTTAACTTAAACCTCGACATTAA	86385		
Qy 627	gcaa 630			
Db 86384	GCTA 86381			
RESULT 15	ATAC011620 100887 Dp	DNA	PLN	18-JAN-2000
LOCUS	Arabidopsis thaliana chromosome III BAC F18C1 genomic sequence,			
DEFINITION	Complete sequence.			
ACCESSION	AC011620			
VERSION	AC011620.6	GI:6714437		
KEYWORDS	HTG.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core			
	eudicots; rosids; eurosids II; Brassicales; Brassicaceae;			
	Arabidopsids.			
REFERENCE	1 (bases 1 to 100887)			
AUTHORS	Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B., Wu,D.,			
	Romling,C.M., Koo,H., Fujii,C.Y., Uterback,T.R., Barnstead,M.E.,			
	Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.			
TITLE	Arabidopsis thaliana chromosome III BAC F18C1 genomic sequence			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 100887)			
AUTHORS	Lin,X. and Kaul,S.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-OCT-1999) The Institute for Genomic Research, 9712			
REFERENCE	Medical Center Dr. Rockville, MD 20850, USA, xlin@tigr.org			
AUTHORS	3 (bases 1 to 100887)			
TITLE	Lin,X.			
JOURNAL	Direct Submission			
COMMENT	Submitted (18-JAN-2000) The Institute for Genomic Research, 9712			
	Medical Center Dr., Rockville, MD 20850, USA			
	On Jan 18, 2000 this sequence version replaced gi:6175135.			
	Address all correspondence to:atetlgr.org			
	BAC clone F18C1 is from Arabidopsis chromosome III and is near the			
	molecular marker mil72.			
	The orientation of the sequence is from SP6 to T7 end of the BAC			
	clone.			
	Genes were identified by a combination of three methods: Gene			
	prediction programs including GRAIL (available by anonymous ftp			
	from athurh.epm.ornl.gov), GeneFinder (Phil Green, University of			
	Washington), Genscan (Chris Burge,			



<http://www.stanford.edu/~chris/GENSCANW.html>), and NetPlantGene (<http://www.cbs.dtu.dk/netgene/cbsnetgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tdb/at.html>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown proteins'. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

## FEATURES

## source

Location/Qualifiers

1. 100887

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="III"

/map="M1172"

/clone="F18C1"

/complement(1..12001)

/note="overlap with BAC clone F10A16 (AC012393:1..12001)."

2971..3040

/rpt\_family="(GAA)n"

/complement(5045..5109)

/rpt\_family="(GAA)n"

/complement(14590..14554)

/rpt\_family="Poly A"

join(<15397..15585,16169..16247,16449..16543,16641..16772,17050..17235,17680..17843,17919..17997,18184..18312,18413..18540,18736..19141)

/gene="F18C1.1"

<15397..>19141

/gene="F18C1.1"

/note="similar to GB:AA027644"

join(15397..15585,16169..16247,16449..16543,16641..16772,17050..17235,17680..17843,17919..17997,18184..18312,18413..18540,18736..19141)

/gene="F18C1.1"

/codon\_start=1

/product="putative importin alpha"

/protein\_id="AAF26125.1"

/db\_xref="GI:6714438"

/translation="MKSGTMSVRSRGYKAVYDVGRRRRRDEVEIRKAKRRESLKRRREALPHSPASDLDOKLISCIWSDERDLLEATQIRTLGCEMNVREVEIQA

GLVRFVEFLTWDSPOLOPEAAMALTNJASGSETEVYIDHGAVALIVRLNSPYD

VVREOYVMAIGNSDSPRCRDYVIGHAALPSILLQINRGAKLSLVNAAMTSLICR

GKQPPEDYSALRALQILRLDDKELATYGCALYVTSQSNKRTQAVIEANVCAR

LIGLSHRSFVTPALRTIGNLVGNDSTQHTIDQLPCLVNLGSGTKTKTRKE

ACWVTSIHRSFQSOQIAVFADICPALVNLQNSGDYKKAAMALINAJAGSGYKO

IMFLVKECIKPLCDLTCSDQLVNVCLAEAKIKLVGEVFSRHAGIYOCPTNV

NPAQLIEEAELEKIEGLQSHENNIDYETAVKILLETYMEEEEDDEQDMITFPV

DNFANMPTSGTSEMHGCP"

19025..19059

/rpt\_family="(GAA)n"

/complement(1011196532..20011,20101..20181,20264..20329,20424..20501,20581..20652,20874..20963,21057..21333,21648..>21871)

/gene="F18C1.2"

/complement(19632..>21871)

/gene="F18C1.2"

/note="similar to putative syntaxin protein GB:CAB52175

from [Arabidopsis thaliana]"

complement(join(19904..20011,20101..20181,20264..20329,20424..20501,20581..20652,20874..20963,21057..21333,21648..>21871))

/gene="F18C1.2"

/codon\_start=1

/product="putative syntaxin protein, AtSNAP33"

/protein\_id="AAF26126.1"

/gene="F18C1.2"

/codon\_start=1

/product="putative syntaxin protein, AtSNAP33"

/protein\_id="AAF26126.1"

/gene="F18C1.2"

/codon\_start=1

/product="putative syntaxin protein, AtSNAP33"

/protein\_id="AAF26126.1"

/gene="F18C1.2"

/codon\_start=1

/product="putative syntaxin protein, AtSNAP33"

/protein\_id="AAF26126.1"

/gene="F18C1.2"

/codon\_start=1

/product="putative syntaxin protein, AtSNAP33"

/db\_xref="GI:6714439"

/translation="WATNRKRLRRKYNLSRVAPMGSSSSSTLEHNSLGANSK

LQPVEMASTNPNRYAPVSTEDPENGSRGTYTGLPDWDVSEETSVYIQAR

TKVAELGKAHALMPNSGDCGEKHQETITQVTFGLKSEKQORLSAAPSIDS

NRKVNQSLADLQNLSELRKQSTYKLRQKDGADLENNLSGKRYKADDFE

DMVNSEHMQSKIRKSEIETREKREIQGVESSEFLAQMKLALVITQGTVDRI

DYNIQVASTYDDGLKQOKAEFRQGGKMCASVILICFIMVLITLKEILL"

22133..22223

/rpt\_family="(TAA)n"

/complement(22179..22226)

/rpt\_family="Poly A"

join(<22810..22885,23454..23596,23884..23922,24150..24328,24593..>24701)

/gene="F18C1.3"

<22810..>24701

/gene="F18C1.3"

join(22810..22885,23454..23596,23884..23922,24150..24328,24593..>24701)

/gene="F18C1.3"

/note="unknown protein"

/codon\_start=1

/protein\_id="AAF26127.1"

/db\_xref="GI:6714440"

/translation="MDSWSDRLASATRRYQLAFPSRSDTFLGFEEDIEEERERF

ACPCSDYFDIYSLCHIDEDHPBAKGVPCVAGVDMSLFGSSCIVSSSS

NVAADPLSEFSPIDAEFTTESCISAEQPVAKTITQCLPEQNAKKTLSAEDHQ

KIKRSEFVEELISRTILDDSL"

complement(27249..27313)

/rpt\_family="(TAA)n"

join(28290..28335,28744..29013,29330..29424,29536..29754,29843..30597)

/gene="F18C1.4"

28290..30597

/gene="F18C1.4"

/note="similar to GB:CAAT4049"

join(28809..29013,29330..29424,29536..29754,29843..30211)

/gene="F18C1.4"

/codon\_start=1

/product="putative transcription factor"

/protein\_id="AAF26128.1"

/db\_xref="GI:6714441"

/translation="MAMQTVRGELFSAPQTSMTWAFSGOPAPESLGDSDSPAGYV

GSYGERGQVYDQSNATHLAFSLGIVASPTLVKPRHATYSQSPCLLEGFSPPTV

IKTFGEQDTGVVSATGSGQSKVMLPMTETEDSTIVNSKQYHGLTRKQSRKAA

VLQKRLSSRCRRKPYMHSHRLHARRRGSGGRFLNKSQNLSSGTNAKKGDSMQ

IOSQRPQOSNSQSEVHHPENGNTQNLNGLNVSQSEVTSMKYLSSPVSHSGMVP

SKWIAAAMNDGNCNFT"

join(36087..36254,36777..37418,37507..37875,37965..38069,38171..38353,38529..38670,38758..38939,39026..39111,39212..39260,39735..39801,39900..40385,40434..40540,41173..41240,41336..41552,41793..42008,42111..42286,42340..42393,42479..42687,42782..42845,43070..43389,43742..43925,44024..46030,46087..46203,46311..46581)

/gene="F18C1.5"

36087..46581

/gene="F18C1.5"

/note="unknown protein"

/codon\_start=1

/protein\_id="AAF26129.1"

/db\_xref="GI:6714442"

/translation="MYRYLNDNIFVRATSPSPALAEVYFVCEGSEKFRKLCNPFYT

PSAPYLEEVAVTNHLVVRGSRSLSLIVGNYIKDQGINILIEERSTYDIYSSE

GNLEDDPLVLSVNRTIEELSSLDIVLPLAAVDPEVYRRLQILKIKEDLATND

VYKFEVDVSGSVSVTDNDVDFLKNNGCAVNSLSDSGLFHDIVRVEDIDTANE

IOESDVALGLEFSEETLYATNSQOLYVMSLPYIQFEDSDCTYLPRLSGKATILDL

SLAFLLSGREGLQPVNSGMDQLYVTFGHGQNSTITITLLGLGVEQATRRSGCE











Query Match	94.2%	Score 924	DB 1	Length 1106
Db	361	TGCCCTGAGCTTGSGAAGCTTGAACTTTACAGTACTCTAGACTCTACAAAACAACT	420	
Qy	421	ccaaggaactataccctccgaacttggaaatctgaagaatctcatcaagcttgatctgta	480	
Db	421	CCAAGGAACCTTACCTCCGAACTTGGAAATCTGAAATCTCATCAGCTTGATCTGTA	480	
Qy	481	caaaaacatcttcagggaatgaatctccactctcttgggaanaattgaagctctgtctt	540	
Db	481	CACAAACAATTTACAGGGAATGATCTCCACTCTCTTTGGAAAATTGAAGCTCTGCTGTT	540	
Qy	541	ttacaggttaatgaacaaccgatitgacccggtlccaatccctagacatccagcgaatccc	600	
Db	541	TTTACGGCTTATATGACAACCATGTAGACCGGTCACATCCCTTAGAGACATCCAGGCAATCCC	600	
Qy	601	aagccttaagttgttgacgtcttaagcaatgatitgtgtggaacaatcccaacaagcg	660	
Db	601	AAGCTTTAAAGTTGTGTGACGTCCTAACCAATGATTTGTGTGAAACAATCCCAACAAACG	660	
Qy	661	acccttctcacatctcccttacgaacttggagaacaaccgagatitggagagaccgga	720	
Db	661	ACCCTTCTCACATTCCTCTTACAGAACTTTGAGAACAACCCGAGATTTGAGGAGCACCGA	720	
Qy	721	attactcgtcttcgaagctcagacactaactgcacttgaaacaactgccaacactgaa	780	
Db	721	ATTACTCGGCTCTTCACACCTACGACACTAATCTGCAACCTGAAACAACTCGAA	780	
Qy	781	aatgaagaatcgggggggagaccttgtaagaacacttccacccttacaatatcaatc	840	
Db	781	AATGAAGAATTTGGGGGGGAGACCTTGTAAGAACACTTCCACACTTATCAAAATTCACATC	840	
Qy	841	tattatcgaataagatatatatatcgttagtaaaaaaacaataaataatgaagaatcgta	900	
Db	841	TATTATGTAAATAGATATATATATGTAGTAAAAAACAATAAATGAAACATCGAATCGGTA	900	
Qy	901	atatcatctgcttccaatttggagaacttcggagctcgtgatgaanaatttccaatgcgatt	960	
Db	901	ATATATATCTGGCTCTCAATTGGAACACTTCGAGGCTGTGTAATTAATTTCTAAATGCGATT	960	
Qy	961	ttcgacctaaattactcacact 981		
Db	961	TTCGCCTAAATTAATCTCACACT 981		
RESULT 2				
ID	V06586	V06586 standard; cDNA to mRNA; 1106 BP.		
AC	V06586;			
DT	03-MUG-1998	(first entry)		
DE	Arabidopsis thaliana SERK LRR homologous EST clone.			
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;			
OS	Arabidopsis thaliana.			
FN	key	Location/Qualifiers		
FT	CDS	142..798		
FT		/*tag= A		
FT		/note= "shows high homology to SERK"		
PN	MO9743427-A1.			
PD	20-NOV-1997.			
PF	13-MAY-1997; E02443.			
PR	14-MAY-1996; GB-010044.			
PI	(NOVS) NOVARTIS AG.			
PA	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;			
DR	WPI. 98-086529/08.			
DR	P-ESDB: W47018.			
PT	Production of apomictic seeds - useful in plant breeding			
PS	Claim 28; Pages 71-73; 123p; English.			
CC	The sequence is that of an EST clone showing high homology to			
CC	SERK LRR (leucine-rich repeat) sequences.			
SO	Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;			

[illegible]



Y065590	ID	V065590	standard; cDNA to mRNA; 1063 bp.
AC	AC	V065590;	
DT	03-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK LRR homologous EST clone.		
KW	receptor kinases; apomixis; apomictic; seeds; production; embryos;		
KM	plant breeding; leucine-rich repeat; ss.		
OS	Arabidopsis thaliana.		
FS	Key	Location/Qualifiers	
FT	CDS	106..762	
FT		/*tag=	
FT		a	
		/note="shows high homology to SERK"	
PN			
PD	20-NOV-1997.		
PF	13-MAY-1997;	E02443.	
PR	14-MAY-1996;	GB-010044.	
PA	(NOVS ) NOVARTIS AG.		
PI	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;		
PI	WPI: 98-086529/08.		
DR	P-PSDB: W47022.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 28; Pages 86-88; 123pp; English.		
CC	The sequence is that of an EST clone showing high homology to		
CC	SERK LRR (leucine-rich repeat) sequences.		
CC	Sequence 1063 bp; 313 A; 242 C; 206 G; 302 T;		

Query Match	85.8%	Score 84.18	DB 1	Length 1063
Seq. Local Similarity	96.1%	Pred. No. 8.9e-224		
Matches 885	Conservative 0	Mismatches 32	Indels 4	Gaps 2
QY	57	ctcaaatctcttcgattccctctcttaaacctccgaaagctcaatgcygclctogaa	116	
Db	59	CTCAAAATCCTAATTAATTAATCTACTCTCTCTCGACCTCCCATAGCTACATAGGCGCTCGAA	118	
OY	117	actatcggtggagagccttcgcagccttgtrtaaccttaactgtgctttgatcaactg	176	
Db	119	ACTATCGGTGGAGACTTTCGACCTTCGTATATCTTAACCTTAGCTTTGATTCAGCTGG	178	
OY	177	tcgaagcaaacctcgaagaagabtcctctctacgctcttcgcgcgagatttcagacatccag	236	
Db	179	TCGAGCAAACTCCGAAGGAGATGCTTTAGCGCTTCGCCGGAGATTAAAGATCCGG	238	
OY	237	accatgcccacagagctggatgccaaacctgttaacctgttaactgtrctatgctatg	296	
Db	239	ACCATGTTCTCCAGAGCTGGAGTCCAACTCTGTAAATCCTTGAACTCGGTTCATGTCA	298	
OY	297	cctgtaaccaagaacacccgcgctactcgtgtgattgttggaaattcaaacactctctgac	356	
Db	299	CCTGTAAACCAAGAACACCCGCGTACTCGTGTGATTTGGGAAATTCAAACTCTCTGTGAC	358	
OY	357	atcttgccctctgagcttggagaagcttggaaacatttcagatctgtagctctcaaaaaa	416	
Db	359	ATCTTGCGCTGAGCTTGGAGACTTGAACATTTCAGATCTGAGCTCTCAAAAACA	418	
OY	417	acatccaaaggaactatacctctcgaacttggaaacttgaagaatctcatcagcttgatc	476	
Db	419	ACATCCAAAGGAATATACCTCTCGAACTTGGAAATCTGAAGATCTCATCAGTTGGATC	478	
OY	477	tgtaacaacaacatcttaacaggaatgtttcccaacttcttttggaaaatttgaagctctg	536	
Db	479	TGTACAAACAACATCTTAAGGAGATGTTCCTCCACTCTTTGGAAAATTAAGTCTCTGG	538	
OY	537	tctttttaagcgttaabgaacaacggaattgacccggtccaatctccttagacatccacgca	596	
Db	539	TCTTTTACGGCTTAATGAACAACCGATGTAGCCGGGCCAATCCCTTAGACATCACTCA	598	
OY	597	tccaaagcccttaagatgttgagcttccaagaatgattttgttggaaaacttcccaaca	656	
Db	599	TCCAAAGCTTAAAGTGTGTGATCTCAACAATGATTTGTGTGAACAATCCCAACA	658	
OY	657	acggaccttgcctacatctcttcaagaacttggaaaacaaccggaatttggaggag	716	

Db	659	ACGACCTTTTGGCTCACACTTCCTTTACAGAACTTTGAGAACCAACCCAGAGTTGGAGGGAC	718
Qy	717	cgggaattactcggctctctgcaagactgacacactaactgcacacttgaaacaactgccaacac	776
Db	719	CGGAATTACTCGAGCTTTTGCAGGCTACAGCACTAACTGCACCTGGAAAAATTGGCAAAACC	778
Qy	777	tgaaaatgaaagaattcgggggggtgacacctgtcaagaacacttcacacactttatcaaatca	836
Db	779	TGAAATGAAAGAAATGGGGGGGTGACCTTTGTAAGACACTTCACCACTTTATCAATATCA	838
Qy	837	cacttatattgtcaacaagtatcatatatagctgtgtaaaaaaacaataatgaagaatcgaatc	886
Db	839	CATCTCTATGTAATPAGATATATATATAGTGTCCAA--AAAAAAATGAAGAATCGAATC	886
Qy	887	ggtataatcaactcgtgtccaaattgaagaacttcgaggtc--tgtatgtaaaaatttctaatt	954
Db	897	AGTAATATCATCTGGGGGTCAATGTGSAACCTTGGAGGTGCTGTGTATGTAATTTCTMAAT	956
Qy	955	ggcgaatttgcgcctaataact 975	
Db	957	GCAGCTTTCGCGTACTGTAAT 977	

ID	Accession	Gene	Location/Qualifiers
AC	V05588	standard: cDNA to mRNA; 788 BP.	
DT	U3-AUG-1998	(first entry)	
DE	Arabidopsis thaliana SERK LRR homologous EST clone.		
KW	receptor kinases; apomictis; apomictic seeds; production; embryos;		
OS	plant breeding; leucine-rich repeat; ss.		
FT	Arabidopsis thaliana.		
FT	Key	Location/Qualifiers	
FT	CDS	2..664	
FT		/*tag= a	
PD	W09743427-A1.	/note= "shows high homology to SERK"	
PF	20-NOV-1997.		
PR	13-MAY-1997; E02443.		
PA	14-MAY-1996; GB-010044.		
PI	(NOVS ) NOVARTIS AG.		
DR	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;		
DR	WPI: 98-085529/08.		
DR	P-FSDB; W47020.		
PT	Production of apomictic seeds - useful in plant breeding		
PS	Claim 28; Pages 79-80; 123pp; English.		
CC	The sequence is that of an EST clone showing high homology to		
CC	SERK LRR (leucine-rich repeat) sequences.		
SO	Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;		



[illegible]

Db	15	CTGGAAAGGGGAGACCCTCTTACACGGGGCTGC	CGGAGCTTATCAGATCCAGACAATGTTGT	74
Qy	247	ccagaagctgggataccaaactcttgtaatactcttgtaacccgtgtccatgtaactgtactgtactgtacaa		306
Db	75	TCAAGATGGGATCCAACTCTGTGTAATCCCTTACTTGGTTCATGATCAGTACCTGTGATCA		134
Qy	307	agacaaccgcgtcactcgltgtgatttgggaaattcaaacctctcctgtgacacttggcgcc		366
Db	135	ACACCATCAATCTACTGCTCGATTTGGGGAAATTCAACTTATCTGGAACATCTAGTACC		194
Qy	367	tgaagctgggaaagcttgaacattcaagtactc-----		400
Db	195	TGAACCTGGGGAAGCTTGAAACTTACATATCTGTATGAAATCATCCTTTTGCTTT		254
Qy	400	-----		400
Db	255	TGATATATCGAAAAACATTTCATTATACGTACACACATATTAACATTTTGGCTTGGACATAT		314
Qy	400	--agagctctacaanaaacaacatccaaaggaaactataccttccgaacttggaaattctga		456
Db	315	TAGTAACTCTACAAAAACGAGATTTCAGGAACTATCTACCTTCTGACCTTGGAAATCTGAA		374
Qy	457	gaattctctcaagcttgatctgtgtcaacaacaactcttaagaggatagttccacctctt		516
Db	375	GAGCTTAATCAGTTGGATGTCGTACAAACAAATCTACCGGGAAATCCATCTTCTT		434
Qy	517	gggaaattgaagctctcgtgctcttlltaagcgttaatgacaacogattgacgggtccaat		576
Db	435	GGGAAATTTGAAGTCACATCTGTTTTTTTTTGGCGCTTAAGAAACCGATGTACCGGCTCTAT		494
Qy	577	ccctgagcaactcaacgcgcaatcccaagccctaaagtgtgaagcttcaagaaagatatt		636
Db	495	TCCTTGAGAACCTCCACAGTTATTATTAACCTTTAAAGTGTGATGATCTCAAGGATGATTT		554
Qy	637	gtgtgaacaatcccaacaacagacccttgcctcacatctcttacaacttgaagaacttggaa		696
Db	555	GTGTGGAACAATTCACATGATGAAAGGACCTTTGAAACACATTCATATGCAAAAATTGAGAA		614
Qy	697	caaccggagatttggaggaccggaattactcgtgctctgcaagctagaacactaacgcac		756
Db	615	CAACTGTGATTTGGGGGACCAAGAACTACTGTGCTTGGCGACTATACACCAATGTGCAC		674
Qy	757	ctggaacaactgccaacccctggaanaatgaaagtgggggttgacccttgaagaacact		816
Db	675	TTAAAAAGAAAGTGAAGAACTATTAAGAGAA-TGTTAGTGACCTGTGAAGAACTGTG		733
Qy	817	caccacttatccaataatc	835	
Db	734	TACCAAGTGTGTGTAATC	752	
RESULT 6				
Db	V06591	V06591 standard; cDNA to mRNA; 2089 BP.		
AC	V06591;			
DT	03-AUG-1998	(first entry)		
DE	Arabidopsis thaliana SERK gene.			
KM	receptor kinase; apomixis; apomictic; seeds; production; embryos;			
KW	Plant breeding; ds.			
OS	Arabidopsis thaliana.			
FH	Key	Location/Qualifiers		
FT	CDS	.195..2072		
FT		/tag= a		
FT		/product= SERK protein		
PN	MO9743427-A1.			
PD	20-NOV-1997.			
PR	13-MAY-1997; E02443.			
PF	14-MAY-1996; GB-010044.			
PI	(NOVS ) NOVARTIS AG.			
PI	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;			
DR	WPI: 98-086529/08.			
PT	P-PSDB; W47023.			
Production of apomictic seeds - useful in plant breeding				







DT 17-JUN-1999 (first entry)  
 DE O. longistaminata Xa21 gene family member A2 DNA.  
 KW Xa21; receptor kinase-like protein; multigene family; RRR; rice;  
 plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
 OS Oryza longistaminata.  
 PN WO9909151-A2.  
 PD 25-FEB-1998.  
 PF 17-JUL-1998; U14841.  
 PR 13-AUG-1997; US-910386.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,  
 PI Wang G;  
 PI MPI: 99-204431/17.  
 PT New RRR polynucleotides and nucleic acid constructs - used for  
 PT generating transgenic plants resistant to Xanthomonas  
 PS Claim 1; Page 52-53; 67pp; English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the Oryza longistaminata and Oryza sativa receptor  
 CC kinase-like protein (RRR) Xa21 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
 CC rice or tomato.  
 SQ Sequence 5940 BP; 1570 A; 1200 C; 1188 G; 1982 T;

Query Match 7.6%; Score 75; DB 1; Length 5940;  
 Best Local Similarity 52.4%; Pred. No. 3.3e-11;

Matches 165; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 341 tcaacctctctgcatcttgccctgagcttggaagcttggaacattatgatactca 400  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 2406 TCGGACTGTCGGGATCATCTCGCGTGGGCAACCTTCTCTCTCGAGACGCTG 2465  
 QY 401 gagctctcaaaaaacaacatccaaggaactatacctccgaacttggaattcgaagaat 460  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 2466 GACCTAGCGCACACCACTGTCCGGCAAGATACCTAGGAAGCTGACGACTCAGCAGG 2525  
 QY 461 ctatcagcttgatctgtacacacaacatctacagggatagttccactctcttgga 520  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 2526 CTCCAACACTGTGATGATTTCAACACGCTATCGGATGAGATTCACGCTCTTGAGGC 2585  
 QY 521 aaattgaagctctgctcttttcttaagcgttaatgaacacgaattgacggtccaatccct 580  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 2586 AACTTAACGACTCTCGGCTTTGAGCTGACTAACAATACACTGCTGAGACCAATCCCT 2645  
 QY 581 agagcactcagcgaatcccaagccttaagttgttgaagcttcaagcaatgatttgt 640  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 2646 TCATCTCTGGGCAAACTCACACGCTCATGATCTTGACATGGCTGAATAATATGCTGTCT 2705  
 QY 641 ggaacaatcccaaca 655  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 2706 AGTTCCATCCCTTCA 2720

RESULT 9

V06570 standard: DNA; 6695 BP.

AC V06570;  
 DT 03-AUG-1998 (first entry)  
 DE Daucus carota SERK gene.  
 KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 OS Daucus carota.  
 FH Key Location/Qualifiers  
 FT CDS 3696..6620  
 FT /\*tag= b  
 FT /note= "contains introns"  
 FT 3731..3802  
 FT /\*tag= b  
 FT /number= 1  
 FT 3851..3979  
 FT /\*tag= c  
 FT Introns

FT /number= 2  
 FT 4124..4211  
 FT /\*tag= d  
 FT /number= 3  
 FT 4284..4357  
 FT /\*tag= e  
 FT /number= 4  
 FT 4430..4528  
 FT /\*tag= f  
 FT /number= 5  
 FT 4642..4757  
 FT /\*tag= g  
 FT /number= 6  
 FT 4890..4967  
 FT /\*tag= h  
 FT /number= 7  
 FT 5295..5803  
 FT /\*tag= i  
 FT /number= 8  
 FT 6197..6339  
 FT /\*tag= j  
 FT /number= 9  
 PN WO9743427-A1.  
 PD 20-NOV-1997.  
 PE 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS ) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 PI MPI: 98-086529/08.  
 DR P-SDB; W47013.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 21; Pages 40-46; 123pp; English.  
 CC The sequence is that encoding SERK, a putative receptor kinase.  
 CC It may be used as part of a method of producing apomictic seeds  
 CC comprising: (a) transforming plant material with a nucleotide  
 CC sequence encoding a protein which in active form in a cell or  
 CC cell membrane renders the cell embryogenic; (b) regenerating  
 CC the transformed material into plants or carpel-containing  
 CC plant parts; and (c) expressing the sequence in the vicinity  
 CC of the embryo sac. The apomictic seeds and embryos thus produced  
 CC can be developed into plant progeny. This is useful in plant  
 CC breeding programs. Controllable and reproducible apomixis provides  
 CC many advantages in plant improvement and cultivar development in  
 CC the case that sexual plants are available as crosses with the  
 CC apomictic plant. Apomixis provides for true-breeding, seed  
 CC propagated hybrids and could shorten and simplify the breeding  
 CC process so that selfing and progeny testing to produce and/or  
 CC stabilise a desirable gene combination could be eliminated.  
 CC Apomixis allows plant breeders to develop cultivars with  
 CC specific stable traits for such characteristics as height,  
 CC seed and forage quality and maturity.  
 SQ Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match

Best Local Similarity 71.0%; Score 74; DB 1; Length 6695;  
 Pred. No. 6.6e-11;

Matches 98; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 193 agagatgctctctgcatcttgccctgagcttggaagcttggaacattatgatactca 252  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 956 AGCGATGCACTTACACAACTTACGAACTTACGAAAGATCCCAACAAATGCTCTGAGAG 1015  
 QY 253 ctggagatcaactctgttaatccttgtaacctggttccatgtaacctgtaacgaacaa 312  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 1016 CTGGGATCCAAACCTTGTGAACCTTGACATGTTTCATGTGACATGAACATGAATAA 1075  
 QY 313 ccgcgtcaactcgtgtgga 330  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 DB 1076 CAGGTGTTAAGACTGA 1093

RESULT 10

T62124







```

Qy      313 ccgcgtcactcgtgtga      330
          | | | | | | | | |
Db      1914 CAGTGTcATAGAGTGTa      1931

```

RESULT	12
X23525	
ID	X23525 standard; DNA; 9424 BP.

DT 17-JUN-1999 (first entry)  
DE O. sativa Xa21 gene family member E DNA.  
KW Xa21; receptor kinase-like protein; multigene family; RRF; rice;  
KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss

PR New RRKpolynucleotides and nucleic acid constructs - used for  
 PT generating transgenic plants resistant to Xanthomonas  
 PS Claim 1: Page 48-50: 67pp: English.  
 CC This invention describes a method for conferring disease resistance in  
 CC plants. The invention describes the use of novel genes and proteins  
 CC belonging to the Oryza longistylamata and Oryza sativa receptor  
 CC kinase like protein (RRK) xaz1 multigene family. Such genes from  
 CC cassava, maize and tomato are also described. The genes and proteins can  
 CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
 CC rice or tomato.  
 SQ Sequence 9424 BP: 2681 A; 2138 C; 2053 G; 2552 T;

Query Match	7.3%	Score 71.8	DB 1	Length 9424
Best Local Similarity	51.7%	Pred. No. 3.1e-10		
Matches 163; Conservative	0	Mismatches 152	Indels 0	Gaps 0

[illegible]

RESULT 13  
T06307  
ID T06307 standard; cDNA; 3573 BP.  
AC T06307.  
DT 14-APR-1996 (first entry)  
DE Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.  
KW Pathogen resistant; Cf-2.2; tomato; C.tulvum; AVR 4; AVR 9; fungal.

KW leaf mould; variegation; ss.  
OS Lycopersicon esculentum.  
PN W09531564-A2.

PF 11-MAY-1995; G01075.  
PR 11-MAY-1994; GB-009394.  
PR 23-DEC-1994; WG-G02817.  
PR 31-MAR-1995; GB-006658.  
PR 07-APR-1995; GB-007233.  
PA (GATS)-GATSBY CHARITABLE FUND.  
PI Hammond-Gosch KE, Jones DA, Jones JDG,  
DR WPI: 96-010949/01.  
P-PSNR: R85299.

PT increasing plant pathogen resistance by induction of variegation -  
 PT may lead to acquired resistance to a broad range of pathogens.  
 PS Claim 9; Page 85-87; 131pp; English.  
 CC 706307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone.  
 CC In a new method this gene is expressed highly in genetic constructs  
 CC which may be used to impart a broad range of pathogen resistance, by  
 CC induction of variegation, to transgenic plants (or parts or propagules  
 CC of plants) containing such constructs. Cf-2.2 imparts resistance to the  
 CC disease caused by the leaf mould fungal pathogen *Cladosporium fulvum*.  
 CC Cf.vulm contains avirulence (Avr) genes that confer recognition by  
 CC plants containing Cf genes, leading to the activation of host defence  
 CC mechanisms to attack the disease.  
 SEQ Sequence 3573 BP; 1032 A; 654 C; 664 G; 1223 T;

Query Match	7.0%;	Score 68.4;	DB 1;	Length 3573;
Best Local Similarity	52.9%;	Pred. No. 1.8e-09;		
Matches 147; Conservative	0;	Mismatches 131;	Indels 0;	Gaps 0

Qy	311	aaccgcgcacatcgctgfgagatttgggnaattcaaacctctcttggaaattcttgcgcctgag	570
Db	995	AAAACCTGTCTAGCTGTGAATCTGTAAATAATACACTTTCGGCTCTATTCCTCTTCA	1054
Qy	371	cttgggaagcttgcacatttacagtatctagagctctacaanaaacacataccaagaagact	430
Db	1055	TTGGGGAACTCGAACCAACTGTGTATGTATTATCTTTACAAATATACAGCTTTCTGGCTCT	1114
Qy	431	atactctccgactcttggaatacttgaagaactcatcaagcttgcattctgcacaacaacat	490
Db	1115	ATTCTGCTTCATTTGGGGAACTCGAACCAACTGTGTATGTGTATCTTTACAAATATACAG	1174
Qy	491	cttaeaaggaagaagttcccaactctcttgggnaattgaagttcctgtctcttttaagcgtc	550
Db	1175	CTTTGTGGCTGTATTCCTGCTTCACTTGGGAAATCTGAACAACCTTGTCTAGGTGTATCTC	1234
Qy	551	aatgcacaacgaattgacgcgttccaatccctagaagcact	588
Db	1235	TACAAATACAGCTTTCTGGCTCTATTCCTGGAAGAAAT	1272

RESULT	14
ID	T06306
AC	T06306
DT	14-APR-1996 (first entry)
DE	Tomato pathogen resistance gene Cf-2.1.
KW	Pathogen resistant; Cf*2.1; tomato; C.fulvum; Avr 4; Avr 9; fungal
KW	leaf mould; variegation; ds.
OS	Lycopersicon esculentum.
FH	Key
FT	Location/Qualifiers
FT	1..1676
FT	5'utr
FT	
FT	cds
FT	1677..5014
FT	/*tag= a
FT	/*tag= b
FT	/product= immature_Cf_2.1_protein
FT	signal_peptide 1677..1745
FT	/*tag= c
FT	mat_peptide 1746..5011
FT	/*tag= d
FT	3'utr 5015..6471



FT W0931564-A2. /tag= e  
PD 23-NOV-1995.  
PF 11-MAY-1995; G01075.  
PR 11-MAY-1994; GB-009394.  
PR 23-DEC-1994; WO-602812.  
PR 31-MAR-1995; GB-006658.  
PR 07-APR-1995; GB-007232.  
PA (GATS-) GATSBY CHARITABLE FOUND.  
PI Hammond-Kosack KE, Jones DA, Jones JDG;  
DR WPI: 96-010949/01.  
PT P-PSDB: R85298.  
PT Increasing plant pathogen resistance by induction of variegation -  
PT may lead to acquired resistance to a broad range of pathogens.  
PS Claim 9; Page 80-83; 131pp; English.  
CC This gene is the tomato pathogen resistance gene Cf-2.1. In a new method  
CC this gene is expressed highly in genetic constructs which may be used  
CC to impart a broad range of pathogen resistance, by induction of  
CC variegation, to transgenic plants (or parts or propagules of plants)  
CC containing such constructs. Cf-2.1 imparts resistance to the disease  
CC caused by the leaf mould fungal pathogen Cladosporium fulvum.  
CC C. fulvum contains avirulence (Avr) genes that confer recognition by  
CC plants containing Cf-genes, leading to the activation of host  
CC defence mechanisms to attack the disease.  
SQ Sequence 6471 BP; 2073 A; 1106 C; 1122 G; 2170 T;

Query Match 7.0%; Score 68.4; DB 1; Length 6471;  
Best Local Similarity 52.9%; Pred. No. 2.3e-09;  
Matches 147; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 311 aaccgcgcacgcctgctgagcttgggaattcaaacctctcgagacattcgccctgag 370  
DB 2679 AAAAATTGCTGAGGTGATCTGTTAAATCAATCAGCTTCTGGCTCTATCTCTCA 2738  
QY 371 ctgggaagcttgaacattacagctatcagagctctacaaaaaacattcaaggaact 430  
DB 2739 TTGGGGATCTGAACAACCTGTATGTATCTTAATAATCAACGCTTTCTGGCTCT 2798  
QY 431 ataccctccgaacttggaaatcgaagaatcattcagcttgcagcttgcagacacaaat 490  
DB 2799 ATTCTGCTCTCATTTGGGGAATCTGAACAACCTGTATGTATGTATCAATATATCAG 2858  
QY 491 cttaacaggagtagtcccaattcttgggaaattgaagctcctgctcttttaagcgtt 550  
DB 2859 CTTTCTGGCTCTATCTCTGTTGTTGGGGAATCTGAACAACCTGTCTAGGTTATCTTC 2918  
QY 551 aatgacacgattgacggtccatccctagagcaact 588  
DB 2919 TACAATAATCAGCTTCTGCTCTATCTCTGAGAAAT 2956

RESULT 15  
X23523 X23523 standard; DNA; 8416 BP.  
AC X23523;  
DT 17-JUN-1999 (first entry)  
DE O. longistaminata Xa21 gene family member A1 DNA.  
KW Xa21; receptor kinase-like protein; multigene family; RKK; rice;  
KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
OS Oryza longistaminata.  
PN W09909151-A2.  
PD 25-FEB-1999.  
PE 17-JUL-1998; U14841.  
PR 13-AUG-1997; US-910386.  
PA (REGC) UNIV CALIFORNIA.  
PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,  
PI Wang G;  
DR WPI: 99-204431/17.  
DR P-PSDB: W93596.  
PT New RKK polynucleotides and nucleic acid constructs - used for  
PT generating transgenic plants resistant to Xanthomonas  
PS Claim 1; Page 37-39; 67pp; English.

CC This invention describes a method for conferring disease resistance in  
CC plants. The invention describes the use of novel genes and proteins  
CC belonging to the Oryza longistaminata and Oryza sativa receptor  
CC kinase-like protein (RKK) Xa21 multigene family. Such genes from  
CC cassava, maize and tomato are also described. The genes and proteins can  
CC be used for enhancing resistance to Xanthomonas in a plant, preferably in  
CC rice or tomato.  
SQ Sequence 8416 BP; 2220 A; 1984 C; 1707 G; 2505 T;

Query Match 6.9%; Score 67.6; DB 1; Length 8416;  
Best Local Similarity 51.0%; Pred. No. 4.3e-09;  
Matches 160; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 341 tcaacctcttggacattctgccttgaagcttgggaagcttgaacattacagatccta 400  
DB 5026 TCCAACTCTGACGGGATCTTCGCCATCGCTGGGCAACCTTCTCCAGAGCCCTG 5085  
QY 401 gagctctacaaaacacacatccaaaggaactatcctccgaacttggaaatcgaagaat 460  
DB 5086 CAACCTAGCAACACACACTCTCCGCAAGATACCCCAAGGCTCAGCCGCTCAGCAGG 5145  
QY 461 ctatcagcttggatctgtacaaacaacatttacaaggaatgtccacttcttggga 520  
DB 5146 CTCACAGCAGCTGTACTGATTTCAACAGCCATCGGCTGAGATTCCAGCTGTTGGGC 5205  
QY 521 aaattgaagctctgtctctttaaagcttaatgaaacgattgacgggtccaatccct 580  
DB 5206 AATCTAACAGTCTCTCACTTCTTAGCTGACTTAACAATACAGTGTCTGGTCTATCCCT 5285  
QY 581 agagcactcagggcaatcccaagccttaagttgttgcagctcgaagaatgttgt 640  
DB 5266 TCATCCCTGGGCAAGCTCACCGGCTCTATCTTGACAGGCTGAATAATATGCTGCT 5325  
QY 641 ggaacaatcccaac 654  
DB 5326 GGTTCATCTCTAC 5339

Search completed: June 23, 2000, 22:53:01  
Job time: 40577 sec



—

■  
.  
.  
.

.. ... ..



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:44:52 ; Search time 274.21 Seconds  
(without alignments)  
465,028 Million cell updates/sec

Title: US-09-180-798-24

Perfect score: 981  
Sequence: 1 agtgcgaatgaattactgttg.....tcgcctaattactcaact 981

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptoddata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptoddata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptoddata/2/ina/5C\_COMB.seq:\*  
4: /cgn2\_6/ptoddata/2/ina/5D\_COMB.seq:\*  
5: /cgn2\_6/ptoddata/2/ina/5E\_COMB.seq:\*  
6: /cgn2\_6/ptoddata/2/ina/PCTUS\_COMB.seq:\*  
7: /cgn2\_6/ptoddata/2/ina/backfilest1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	6.8	3921	4 US-08-567-375-3	Sequence 3, Appl1
2	67	6.8	3921	4 US-08-567-680A-3	Sequence 3, Appl1
3	67	6.8	5992	3 US-08-475-891A-3	Sequence 3, Appl1
4	60.2	6.1	2075	1 US-08-238-163-3	Sequence 3, Appl1
5	60.2	6.1	6256	3 US-08-475-891A-1	Sequence 1, Appl1
6	60.2	6.1	6256	4 US-08-567-375-1	Sequence 1, Appl1
7	60.2	6.1	6256	4 US-08-567-680A-1	Sequence 1, Appl1
8	56.8	5.8	1058	1 US-08-238-163-1	Sequence 1, Appl1
9	56.8	5.8	5733	3 US-08-473-553A-1	Sequence 1, Appl1
10	54.2	5.5	1554	4 US-08-587-680A-24	Sequence 24, Appl1
11	49.2	5.0	831	4 US-08-367-375-15	Sequence 15, Appl1
12	48.4	4.9	3905	4 US-08-666-271-4	Sequence 4, Appl1
13	48.4	4.9	3905	4 US-08-666-271-1	Sequence 1, Appl1
14	47.4	4.8	7218	1 US-08-232-663-14	Sequence 14, Appl1
15	41.4	4.2	1116	2 US-08-244-646-14	Sequence 14, Appl1
16	41.4	4.2	2917	2 US-08-592-936B-20	Sequence 20, Appl1
17	41.4	4.2	2917	4 US-08-111-573-20	Sequence 20, Appl1
18	40.8	4.2	792	2 US-08-244-646-16	Sequence 16, Appl1
19	36.4	3.7	457	3 US-08-189-256A-16	Sequence 16, Appl1
20	36.4	3.7	1448	3 US-08-189-256A-11	Sequence 11, Appl1
21	34.6	3.5	2088	3 US-08-602-264A-11	Sequence 11, Appl1
22	34.6	3.5	246240	3 US-08-724-394A-20	Sequence 20, Appl1
23	34.6	3.5	246240	3 US-08-724-394A-21	Sequence 21, Appl1
24	34.6	3.5	246240	3 US-08-724-394A-22	Sequence 22, Appl1
25	34	3.5	251	3 US-08-623-906A-16	Sequence 16, Appl1
26	33.8	3.4	1988	3 US-08-508-786-2	Sequence 2, Appl1
27	33.8	3.4	1988	6 PCT-US96-12158-2	Sequence 2, Appl1

C	28	33.8	3.4	2010	3	US-08-508-786-1	Sequence 1, Appl1
C	29	33.8	3.4	2010	6	PCT-US96-12158-1	Sequence 1, Appl1
C	30	33.6	3.4	2363	1	US-08-096-947-2	Sequence 2, Appl1
C	31	33.6	3.4	2363	2	US-07-919-140B-2	Sequence 2, Appl1
C	32	33.6	3.4	2363	6	PCT-US93-06939-2	Sequence 2, Appl1
C	33	33	3.4	291	2	US-08-442-063A-12	Sequence 3, Appl1
C	34	33	3.4	426	2	US-08-442-063A-35	Sequence 35, Appl1
C	35	33	3.4	567	2	US-08-442-063A-38	Sequence 38, Appl1
C	36	33	3.4	711	2	US-08-442-063A-41	Sequence 41, Appl1
C	37	33	3.4	849	2	US-08-442-063A-44	Sequence 44, Appl1
C	38	33	3.4	924	2	US-08-442-063A-47	Sequence 47, Appl1
C	39	33	3.4	1002	2	US-08-442-063A-26	Sequence 26, Appl1
C	40	33	3.4	1026	1	US-08-272-919-1	Sequence 1, Appl1
C	41	33	3.4	1026	2	US-08-619-916-1	Sequence 1, Appl1
C	42	33	3.4	1026	6	PCT-US95-08342-1	Sequence 1, Appl1
C	43	33	3.4	1593	7	5340934-3	Sequence 1, Appl1
C	44	33	3.4	4673	1	US-07-638-431-1	Sequence 1, Appl1
C	45	33	3.4	4673	6	PCT-US92-00018-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-08-567-375-3  
: Sequence 3, Appl1 US/08567375  
: Patent No. 5952485  
: GENERAL INFORMATION:  
: APPLICANT: Ronald, Pamela C.  
: APPLICANT: Wang, Guo-Liang  
: APPLICANT: Song, Wen-Yuang  
: APPLICANT: Szabo, Veronique  
: TITLE OF INVENTION: Procedures and Materials for Confering  
: TITLE OF INVENTION: Disease Resistance in Plants  
: NUMBER OF SEQUENCES: 16  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Townsend and Townsend and Crew LLP  
: STREET: Two Embarcadero Center, Eighth Floor  
: CITY: San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94111-3834  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/567,375  
: FILING DATE: 04-DEC-1995  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 60/004,645  
: FILING DATE: 29-SEP-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/475,891  
: FILING DATE: 07-JUN-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/373,375  
: FILING DATE: 17-JAN-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Bastian, Kevin L.  
: REGISTRATION NUMBER: 34,774  
: REFERENCE/DOCKET NUMBER: 023070-058930  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 576-0200  
: TELEFAX: (415) 576-0300  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 3921 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear



MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1..2676, 3520..3918)  
OTHER INFORMATION: /product= "Xa-21"  
US-08-567-375-3

Query Match 6.8%; Score 67; DB 4; Length 3921;  
Best Local Similarity 49.9%; Pred. No. 4.7e-10;  
Matches 169; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 344 aacctctcgacatcttgccgctgagcttggaagcttgaacattacagatctcagag 403  
DB 1159 AAGATCAGGAGACATTCGAGAGATATTGGCATTTTATGGCTTACACATCTCTAT 1218  
QY 404 cctacaacaaacacatcccaaggaactacatctcgaacttggaaacttgaagatctc 463  
DB 1219 CTCTGCAACAAATTTCAGAGGGCTCTTCATCATCGTTGGCAGGCTTAAAACTTA 1278  
QY 464 atcagcttgatctgtacaacaaacattcagaagtagtcccaattcttgggaaa 523  
DB 1279 GGCATTTCTACGCGCTCGAAGAAACACTTGAGCGGTTGATCCCGTTGGCCATAGGAAT 1338  
QY 524 ttgaagctctgctcttcttcaagcttaattgacacccgattgacggctcaatccctaga 583  
DB 1339 CTTACTGAACTTAATATCTTACTGCTGGCACCAACAAATTCAGTTGGATACCATAC 1398  
QY 584 gcactacgagcaatcccaagccttaagtgttgacgctcgaagaattgtgtgga 643  
DB 1399 ACACCTCAAACTCACAACAACTTGTTGATTAAGCCCTTCAACTAATACCTTAGTGT 1458  
QY 644 acaatcccaacaaacgagacccttgcacattcctta 682  
DB 1459 CCAATACCCAGTGAATTATTCATTAATTCAAACACTATCA 1497

## RESULT 2

US-08-587-680A-3  
Sequence 3, Application US/08587680A  
Patent No. 5877434  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Conferring  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/587,680A  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,375  
FILING DATE: 17-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/475,891  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,645

FILING DATE: 29-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/567,375  
FILING DATE: 04-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-058940US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1..2676, 3520..3918)  
OTHER INFORMATION: /product= "Xa-21"  
US-08-587-680A-3

Query Match 6.8%; Score 67; DB 4; Length 3921;  
Best Local Similarity 49.9%; Pred. No. 4.7e-10;  
Matches 169; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 344 aacctctcgacatcttgccgctgagcttggaagcttgaacattcagatctcagag 403  
DB 1159 AAGATCAGGAGACATTCGAGAGATATTGGCATTTTATGGCTTACACATCTCTAT 1218  
QY 404 cctacaacaaacacatcccaaggaactacatctcgaacttggaaacttgaagatctc 463  
DB 1219 CTCTGCAACAAATTTCAGAGGGCTCTTCATCATCGTTGGCAGGCTTAAAACTTA 1278  
QY 464 atcagcttgatctgtacaacaaacattcagaagtagtcccaattcttgggaaa 523  
DB 1279 GGCATTTCTACGCGCTCGAAGAAACACTTGAGCGGTTGATCCCGTTGGCCATAGGAAT 1338  
QY 524 ttgaagctctgctcttcttcaagcttaattgacacccgattgacggctcaatccctaga 583  
DB 1339 CTTACTGAACTTAATATCTTACTGCTGGCACCAACAAATTCAGTTGGATACCATAC 1398  
QY 584 gcactacgagcaatcccaagccttaagtgttgacgctcgaagaattgtgtgga 643  
DB 1399 ACACCTCAAACTCACAACAACTTGTTGATTAAGCCCTTCAACTAATACCTTAGTGT 1458  
QY 644 acaatcccaacaaacgagacccttgcacattcctta 682  
DB 1459 CCAATACCCAGTGAATTATTCATTAATTCAAACACTATCA 1497

## RESULT 3

US-08-475-891A-3  
Sequence 3, Application US/08475891A  
Patent No. 5859339  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Conferring  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: us/08/475,891A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,375  
FILING DATE: 17-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 02370-05891005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5992 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(512..3149, 3993..4393)  
OTHER INFORMATION: /product= "RRK-B"  
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease  
OTHER INFORMATION: resistance gene RRK-B from rice (Oryza  
OTHER INFORMATION: sativa)"  
US-08-475-891A-3

Query Match 6.88; Score 67; DB 3; Length 5992;  
Best Local Similarity 49.9%; Pred. No. 5.8e-10;  
Matches 169; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 344 aaccctctggacatctcgccgctgagcttgagagcttgaaacttaagctatcagag 403  
DB 1670 AAGATCAGAGAGAGATTCGAGAGATTTGGCAATCTTATTGGCTTACAACTCTCTAT 1729  
QY 404 ctctcaaaaaaacatccaaaggaactataccttcgaaacttgaaactc 463  
DB 1730 CTCTCAACAACAATTGAGAGGCTCTTCATCATCGTTGGCAGGCTTAAAACTTA 1789  
QY 464 atcagcttgatctgttaacaaacatcttcagagatagttcccaactctttgggaaa 523  
DB 1790 GGCACTTCTACGCGCTACGAAACAACTTGAGCGGTTCGATCCCGTTGGCCATGAGAAAT 1849  
QY 524 ttgaagctctgctcttttaagcttaataacacacgattgacgggtccatccctaga 583  
DB 1850 CTTATGAGAACTTAATATTCTACTGCTCGGCAACAATTCAGTGGATACCATAC 1909  
QY 584 gcactcaaggaacatcccaagccttaagttgtgaagctcaagaatgattgtgtga 643  
DB 1910 ACACCTCTAAACCTCAAACTTGTGTGATGAGCGCTTCACTAATACTTAGTGT 1969  
QY 644 acaatcccaaaaacgagaccttgcctcaacttcctta 682  
DB 1970 CCAATACCCAGTGAATTTATCAATATCAACACTATCA 2008

RESULT 4  
US-08-238-163-3  
Sequence 3, Application US/08238163  
Patent No. 5569830  
GENERAL INFORMATION:  
APPLICANT: BENNETT, Alan  
APPLICANT: LABAYTICH, John M.  
APPLICANT: POWELL, Ann  
APPLICANT: STOLT, Henrik

TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/238,163  
FILING DATE: 03-MAY-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 2307E-540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 421..1401  
US-08-238-163-3

Query Match 6.1%; Score 60.2; DB 1; Length 2075;  
Best Local Similarity 49.0%; Pred. No. 3.3e-08;  
Matches 221; Conservative 0; Mismatches 223; Indels 7; Gaps 2;

QY 204 tctacgctcttcgcgcgagttgacagatccagacatgtctccagagctggatccaa 263  
DB 513 TCTCAAAATAAAGAAAGACTTAGCAATCTTACCAATTTAGCTTCAATGGAGATCCAAAC 572  
QY 264 ctcttgtaaaccttgtaactggttcacatgtcaacctgttaacaaagacccgctac 323  
DB 573 AGATTGCTGTACGTGATGATCAATAAATGTGACCGGAAAC---CAACCGGATAAATG 628  
QY 324 gtgtgatttgggaatccaacacctctgacatcttgaccttgaccttgaggaaactg 383  
DB 629 CTCACACCGCTTCCCAAGCAATATCTCCGCCAAATTCGGGACCGCTCGAGACTTC 688  
QY 384 aacattacagatctcagagctctacaa---aaacaacatccaaaggaactataccttcg 440  
DB 689 CATATCTGGAACATTTGGAATTTTCATCATGTTACTATCTCACCGGAACAAATCTCACCTG 748  
QY 441 aacttggaaacttgaagatctcagctgagctgtgacacaaacaaatcttaacagga 500  
DB 749 CAATTGCGAAGCTCAAAATCTCAAAATGTTAAGCTCAAGCTTCACTTAACCTTACAGTTC 808  
QY 501 tagttcccaactcttgggaatgaagtctcgtgcttttaagcgttaatgaacac 560  
DB 809 CGATCCCTGAATTCCTTAGTACGCTGAAGAAATTTGAGCTGCTCGATGCAATTAACATC 868  
QY 561 gatgaacggtccatccctcagagcactcagcgaatcccaagccttaagttgtgtga 620  
DB 869 AATTACCGGAACATTCCTTCTTCCCTCTCTACAGCTTCGAAATTTGCTTAGCATGACT 928  
QY 621 tctcaagcaatgattgtgtggaacatccc 651







Query Match	6.1%;	Score 60.2;	DB 4;	Length 6256;
Best Local Similarity	51.1%;	Pred. No. 5.6e-08;		
Matches 168; Conservative	0;	Mismatches 158;	Indels 3;	Gaps 1;

QY	327	tgatttgggaattcaaacctcttcggacatcttcggccttgagattggaaagcttcgac	386
Db	2735	TGCACTTGGGAGAAATTAACCTGGGGGAGTCTTCTTAATCGTTTCCAACTCTTTCCA	2794
QY	387	att---tacagatctagagctctacaanaacacatccaaagaaactaaccttcgac	443
Db	2795	CTTCGCTTAATTTTCTTCTGCACTTGATTTTGATTAAGATACAGAAAGCATTTCCAAAGGATA	2854
QY	444	ttagaattcgaagaatcctcatcagctttagatctgacaacaacatcttaagagatag	503
Db	2855	TTCGCAATCTTATTGGTTTCAACATCTCTATCTCTGCACAAACAATTTGCAAGGGTCA	2914
QY	504	ttcccaactctttgggaanaattgaagctctcgtctcttllacggttaatgacaacgat	563
Db	2915	TTCATATATGTTGGGGCAGGCTTAGAAATCTTAAGCATTTTAATCGGCTTACGAAAACAAT	2974
QY	564	tgaacggtccaatccccaagagcaactacgycgaatcccaagcccttaaaagtgtgacgct	623
Db	2975	TGACGGTTCGATGCCCAATTGGCATTAGGAATCTTAAGTAATATCTTACTGTCG	3034
QY	624	caagcaatgatattgctgtggaacaatccca	652
Db	3035	GCACCAACAATTTAGTGGTTGGATACCA	3063

RESULT 7  
 US-08-587-680A-1  
 Sequence 1, Application US/08587680A  
 Patent No. 5977434  
 GENERAL INFORMATION:  
 APPLICANT: Ronald, Pamela C.  
 APPLICANT: Wang, Guo-Liang  
 APPLICANT: Song, Wen-Yuang  
 APPLICANT: Szabo, Veronique  
 TITLE OF INVENTION: Procedures and Materials for Conferring  
 TITLE OF INVENTION: Disease Resistance in Plants  
 NUMBER OF SEQUENCES: 27  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentln Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/587,680A  
 FILING DATE: 17-JAN-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/373,375  
 FILING DATE: 17-JAN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/475,891  
 FILING DATE: 07-JUN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/004,645  
 FILING DATE: 29-SEP-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/567,375  
 FILING DATE: 04-DEC-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774

```

1  REFERENCE/DOCKET NUMBER: 023070-05894005
2  TELECOMMUNICATION INFORMATION:
3  TELEPHONE: (415) 576-0200
4  TELEFAX: (415) 576-0300
5  INFORMATION FOR SEQ ID NO: 1:
6  SEQUENCE CHARACTERISTICS:
7  LENGTH: 6256 base pairs
8  TYPE: nucleic acid
9  STRANDEDNESS: single
10 TOPOLOGY: linear
11 MOLECULE TYPE: DNA (genomic)
12 FEATURE:
13 NAME/KEY: CDS
14 LOCATION: join(1648..4383, 5178..5513)
15 OTHER INFORMATION: /product= "RRK-F"
16 OTHER INFORMATION: /note= "Xanthomonas spp. disease
17 OTHER INFORMATION: resistance gene RRR-F from rice (Oryza
18 OTHER INFORMATION: sativa)"
19 JS-08-587-680A-1

```

Query Match	6.1%;	Score 60.2;	DB 4;	Length 6256;
Best Local Similarity	51.1%;	Pred. No. 5.6e-08;		
Matches 169;	Conservative	0;	Mismatches 158;	Indels 3; Gaps 1;

[illegible]

RESULT 8  
 US-08-238-163-1  
 ? Sequence 1, Application US/08238163  
 ? Patent No. 5569830  
 ?  
 ? GENERAL INFORMATION:  
 ?  
 ? APPLICANT: BENNETT, Alan  
 ?  
 ? APPLICANT: LABAVITCH, John M.  
 ?  
 ? APPLICANT: POWELL, Ann  
 ?  
 ? APPLICANT: STOLTZ, Henrik  
 ?  
 ? TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
 ?  
 ? TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE  
 ?  
 ? NUMBER OF SEQUENCES: 24  
 ?  
 ? CORRESPONDENCE ADDRESS:  
 ?  
 ? ADDRESSEE: Townsend and Townsend Kourie and Crew  
 ?  
 ? STREET: Stewart Street Tower, One Market Plaza  
 ?  
 ? CITY: San Francisco  
 ?  
 ? STATE: California  
 ?  
 ? COUNTRY: US  
 ?  
 ? ZIP: 94105-1493  
 ?  
 ? COMPUTER READABLE FORM:  
 ?  
 ? MEDIUM TYPE: Floppy disk  
 ?  
 ? COMPUTER: IBM PC compatible  
 ?  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ?















Db 2913 TATCAATCTCTCAAGACAGATTGTGAGCTATTTCCAGATTAATTGGAGATCTTCT 2972  
QY 457 gaatcacaagcttgatctgtacacaacaactctacaggatgttccactcttt 516  
Db 2973 TGGACTTCGTACGTTGAATCTTGCTCACAATGTCTTGAAGGTCAATATACCGGACATCAT 3032  
QY 517 gggaaattgaagctctgtcttttaagcgttaataagacaacgattgacgggtccatt 576  
Db 3033 TCATAATTATTCAGTACTCAATCTTGGATCTCTCATCTAATAAATACGCGAGAAAT 3092  
QY 577 cctaagagcaactcagcgcaactccaagccttaagattgttgacgtctcaagcaatatt 636  
Db 3093 TCCGACAGAGTTCATCCCTCAATCTCTGAAGTCTTAATCTCTCAACAATCATCT 3152  
QY 637 gtgtggaacaatcccaacaacgagcccttgcct 670  
Db 3153 TGTGTGATGATCCCAAGAAACAAATTTGAT 3186

## RESULT 14

US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHREIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpc-Fls  
US-08-232-463-14

Query Match 4.8%; Score 47.4; DB 1; Length 7218;  
Best Local Similarity 3.5%; Pred. NO. 0.00032;  
Matches 12; Conservative 194; Mismatches 135; Indels 0; Gaps 0;

QY 12 tttagttgcttctcctcctcttcttgaagaatttccttactcctaattcctttc 71  
Db 1115 YY 1174  
QY 72 gattccctctcttaacctcgaagctcaatgagcgctcgaactatcggtggagc 131  
Db 1175 YY 1234  
QY 132 tcttcagacttcgttaaccttaaccttagcttgaattcaactggtcgaagaactcgc 191  
Db 1235 YY 1294  
QY 192 aagagatgctctcctcagctcctcgcgcgaggttgacaatcagaacatgctccca 251  
Db 1295 YY 1354  
QY 252 gctggatccaactctgttaactcctgtactgctcatgtaacctgtaaccaagaca 311  
Db 1355 YY 1414  
QY 312 accgagctcaactcgttgagatttggaattcaacctctcct 352  
Db 1415 YYYYYYYYYYYYYYYYYYGTACCAATCTCTCATCTCT 1455

## RESULT 15

US-08-244-646-14  
Sequence 14, Application US/08244646  
Patent No. 5744692  
GENERAL INFORMATION:  
APPLICANT: Cervone, Felice  
APPLICANT: De Lorenzo, Giulia  
APPLICANT: Salvi, Giovanni  
APPLICANT: Albersheim, Peter  
APPLICANT: Darvill, Alan  
APPLICANT: Bergmann, Carl  
TITLE OF INVENTION: Nucleotide Sequences Coding An  
TITLE OF INVENTION: Endopolygalacturonase Inhibitor  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sally A. Sullivan  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,646  
FILING DATE: 06-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT RM 91A 000915  
FILING DATE: 06-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/IT/00158  
FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Sullivan, Sally A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 19-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)498-8080  
TELEFAX: (303)498-8089  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1116 base pairs  
TYPE: nucleic acid



```
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Phaseolus vulgaris
STRAIN: Saxa
IMMEDIATE SOURCE:
CLONE: lambda PGIP-3.3
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1026
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1027..1116
US-08-244-646-14
```

```
Query Match 4.28; Score 41.4; DB 2; Length 1116;
Best Local Similarity 48.18; Pred. No. 0.0072;
Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 409 caaaacacatccaaggaactatacttcggaacttgaaatcgaagaatctcatcag 468
   || ||||| || || || || || || || || || || || || || || || ||
DB 342 CATCAATTAACCTGTCGGTCCAAATCCCCCGCCATCGCTAACTCACCACACTCCACTA 401

QY 469 ctggatctgtacaacaacaatctacagggatagttccactcttgggaaattgaa 528
   || || || || || || || || || || || || || || || || || || || ||
DB 402 TCTTATATCACTCACACCAATGTCGCGCAATACCCGATTTCTGTCCACAGATCAA 461

QY 529 gctctggtctttaaaggcttaataagaaacgatgacccggtcgaatccctagagcaat 588
   || || || || || || || || || || || || || || || || || || || ||
DB 462 AACCCCTGTCACCTGAGCTTCTCTACACGCCCTCTCCGCACTCCCTCCCTCCAT 521

QY 589 cagggcaatcccaagccttaaaagtgtgaagctctcaagcaatgattgtgtgaaacaat 648
   || || || || || || || || || || || || || || || || || || || ||
DB 522 GTCTTCTCTCCCAACCTCGAGGATCATTCGACGCAACGAATCTCCGGCCCAT 581

QY 649 ccc 651
   |||
DB 582 CCC 584
```

Search completed: June 23, 2000, 22:45:22  
Job time: 40268 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:32 ; Search time 6198.48 Seconds  
(without alignments)  
641.482 Million cell updates/sec

Title: US-09-180-798-24  
Perfect score: 981  
Sequence: 1 agtgtgagtaattagtgtg.....tcgcctaatactacacact 981

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202661650 residues  
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

EST:\*

1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: em\_est10:\*  
11: em\_est11:\*  
12: em\_est12:\*  
13: em\_est13:\*  
14: em\_est14:\*  
15: em\_est15:\*  
16: em\_est16:\*  
17: em\_est17:\*  
18: em\_est18:\*  
19: em\_est19:\*  
20: gb\_est1:\*  
21: gb\_est2:\*  
22: gb\_est3:\*  
23: gb\_est4:\*  
24: gb\_est5:\*  
25: gb\_est6:\*  
26: gb\_est7:\*  
27: gb\_est8:\*  
28: gb\_est9:\*  
29: gb\_est10:\*  
30: gb\_est11:\*  
31: gb\_est12:\*  
32: gb\_est13:\*  
33: gb\_est14:\*  
34: gb\_est15:\*  
35: gb\_est16:\*  
36: gb\_est17:\*  
37: gb\_est18:\*  
38: gb\_est19:\*  
39: gb\_est20:\*  
40: gb\_est21:\*  
41: gb\_est22:\*  
42: gb\_est23:\*  
43: gb\_est24:\*  
44: gb\_est25:\*

45: gb\_est26:\*  
46: gb\_est27:\*  
47: gb\_est28:\*  
48: gb\_est29:\*  
49: gb\_est30:\*  
50: gb\_est31:\*  
51: gb\_est32:\*  
52: em\_est20:\*  
53: em\_est21:\*  
54: em\_est22:\*  
55: em\_est23:\*  
56: em\_est24:\*  
57: em\_est25:\*  
58: em\_est26:\*  
59: gb\_est33:\*  
60: gb\_est34:\*  
61: gb\_est35:\*  
62: gb\_est36:\*  
63: gb\_est37:\*  
64: gb\_est38:\*  
65: em\_est27:\*  
66: em\_est28:\*  
67: em\_est29:\*  
68: em\_est30:\*  
69: gb\_est39:\*  
70: gb\_est40:\*  
71: gb\_est41:\*  
72: gb\_est42:\*  
73: gb\_est43:\*  
74: gb\_est44:\*  
75: em\_est31:\*  
76: em\_est32:\*  
77: em\_est33:\*  
78: em\_est34:\*  
79: gb\_est45:\*  
80: gb\_est46:\*  
81: gb\_est47:\*  
82: gb\_gss1:\*  
83: gb\_gss2:\*  
84: gb\_gss3:\*  
85: gb\_gss4:\*  
86: em\_gss1:\*  
87: em\_gss2:\*  
88: em\_gss3:\*  
89: em\_gss4:\*  
90: gb\_gss5:\*  
91: gb\_gss6:\*  
92: gb\_gss7:\*  
93: gb\_gss8:\*  
94: gb\_gss9:\*  
95: em\_gss5:\*  
96: em\_gss6:\*  
97: em\_gss7:\*  
98: em\_gss8:\*  
99: em\_gss9:\*  
100: em\_gss10:\*  
101: em\_gss11:\*  
102: gb\_gss10:\*  
103: gb\_gss11:\*  
104: em\_gss12:\*  
105: gb\_gss12:\*  
106: gb\_gss13:\*  
107: gb\_gss14:\*  
108: gb\_gss15:\*  
109: gb\_gss16:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query

SUMMARIES



NO.	Score	Match	Length	DB	ID	Description
1	403	41.1	469	63	A1992868	A1992868 701493826
2	384.2	39.2	556	23	R89998	R89998 16353 Lambd
3	359.6	36.7	534	33	AA394359	AA394359 25942 Lam
4	353.2	36.0	447	23	H36800	H36800 14929 Lambd
5	343.6	35.0	401	37	AA712221	AA712221 31949 Lam
6	335.8	34.2	353	23	H37300	H37300 15429 Lambd
7	333	33.9	364	42	A1100683	A1100683 33804 Lam
8	328.8	33.5	336	20	Z34187	Z34187 AT353221 Ve
9	315.6	32.2	348	42	A1100682	A1100682 33803 Lam
10	312.6	31.9	608	81	AM443205	AM443205 EST308135
11	311.6	31.8	720	64	AM031110	AM031110 EST274417
12	305	31.1	337	20	Z34606	Z34606 AT353366 Ve
13	298.2	30.4	339	42	A1100679	A1100679 33800 Lam
14	293.8	29.9	646	74	AM221939	AM221939 EST296750
15	292.8	29.8	657	80	AM350720	AM350720 GM210009A
16	291.8	29.7	336	23	H37296	H37296 15425 Lambd
17	288.8	29.4	511	74	AM221278	AM221278 EST297747
18	286.2	29.2	714	64	AM030188	AM030188 EST273443
19	276.8	28.2	334	23	H37195	H37195 15324 Lambd
20	272.8	27.8	695	64	AM038168	AM038168 EST279825
21	270.4	27.6	465	45	A1352795	A1352795 MB61-10D
22	268	27.3	599	63	AM011134	AM011134 S117B03 P
23	267.4	27.3	599	74	AM219797	AM219797 EST302279
24	267.2	27.2	305	25	N65416	N65416 20456 Lambd
25	260.6	26.5	637	80	AM350549	AM350549 GM210009A
26	256	26.1	336	42	A1100680	A1100680 33801 Lam
27	255.8	26.1	676	74	AM220075	AM220075 EST302558
28	254	26.1	442	42	A1100481	A1100481 34856 Lam
29	252.8	25.8	430	79	AM307218	AM307218 s154c07.Y
30	251.8	25.7	569	59	A1775448	A1775448 EST256548
31	244	24.9	522	45	A1441759	A1441759 s482d08.Y
32	241.6	24.6	360	42	A1100685	A1100685 33806 Lam
33	241.2	24.6	545	64	AM037836	AM037836 EST279465
34	239.4	24.4	667	79	AM185847	AM185847 s60d04.Y
35	231.4	23.6	515	79	AM279515	AM279515 s190e09.Y
36	231.2	23.6	612	47	A1496325	A1496325 sb05c09.Y
37	230.4	23.5	541	64	AM036865	AM036865 614019G10
38	229.6	23.4	619	64	AM040482	AM040482 EST283442
39	227.2	23.2	304	36	T21150	T21150 3158 Lambd
40	224.2	22.9	568	47	A1487272	A1487272 EST245594
41	223.6	22.8	423	35	C22371	C22371 C22371 R1ce
42	221.8	22.6	570	50	A1676939	A1676939 605047A07
43	218.2	22.2	616	42	AF074734	AF074734 AF074734
44	215.8	22.0	540	64	AM033599	AM033599 EST277170
45	214.6	21.9	258	42	A1100678	A1100678 33799 Lam

## ALIGNMENTS

RESULT	1	LOCUS	469 bp	MRNA	EST	08-SEP-1999
A1992868		701493826	A. thaliana	Ohio State clone set	Arabidopsis thaliana	
DEFINITION		CDNA clone 701493826, mRNA sequence.				
ACCESSION		A1992868				
VERSION		A1992868.1	GI:5839773			
KEYWORDS		EST.				
SOURCE		thale cress.				
ORGANISM		Arabidopsis thaliana				
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi				
AUTHORS		1 (bases 1 to 469)				
		Chen, J., Komiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mounoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrta, A., Murry, L., Turner, C., Kikorian, S., Elder, L., and				

TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN
Hanson, D.	Unpublished (1999)	On Dec 20, 1995 this sequence version replaced gi:1135328. Contact: David Smoller, Ph.D. Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc. 4633 World Parkway Circle, St. Louis, MO 63134, USA Tel: 877-577-2733 Fax: 314-427-3324 Email: service@genomesystems.com.	source 1..469 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="701493826" /note="CDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."	120 a 124 c 89 g 136 t	Query Match 41.1%; Score 403; DB 63; Length 469; Best Local Similarity 96.5%; Pred. No. 8,6e-92; Matches 412; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
57	ctcaatctcttgcagcttcctcctccttaaacctcgaagagctcagcgctcga	116			
43	CTCAATTCCTTATGATTAATCTCTCTCTGACCTCGAAGCTCAAGGCTTCGAA	102			
117	actacgctggaagcttcctcagcttcgttaacaccttaaccttagcttgattcaactgg	176			
103	ACTATGCTGGAGCTTCCTGCAAGCTTGCTTAACCTTAGCTTGGATTCACCTGG	162			
177	tccaagcaacctccgaagagatgctctcagctcttcgcggagttggaagatccag	236			
163	TCCAAGCAACCTCCGAAGGAGATGCTTTACCGCTTCGCCGAGTTTAACGATCCG	222			
237	accatgctccccaagcttggaatccaactctgttaaccttgatcgttcacatgca	296			
223	ACCATGCTTCACAGCTGGAGATCCAACTTGTATACCTTGATTCGATCA	282			
297	cctgtaaccaagaacacgcgtcactcgctggtgatttgggaattcaaacctctgcag	356			
283	CCTGTAACCAAGACACCGGCTCCTGCTGATTTGGGGAATTCGAACCTCTCGAC	342			
357	atcttgccctgagcttggaagcttgaacattacagttactaaggtctacaanaaa	416			
343	ATCTTGCCCTGAGCTTGGAGACTTGAACATTACAGTATCTAGAGCTTACAA	402			
417	acatccaagaacatactctcgaacttgaaatctgaaatctcaatcagcttgatc	476			
403	ACATCCAAGAACATATACCTTCCGAATCTGGAATCTGAAGATCATCAGCTTGATC	462			
477	tgtaaca 483				
463	TGTACAA 469				

RESULT	2	LOCUS	556 bp	MRNA	EST	30-DEC-1997
R89998		16353	Lambda-PRI2	Arabidopsis thaliana	cDNA clone 187H577, mRNA	
DEFINITION		sequence.				
ACCESSION		R89998				
VERSION		R89998.1	GI:957538			
KEYWORDS		EST.				
SOURCE		thale cress.				
ORGANISM		Arabidopsis thaliana				
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;				



REFERENCE 1 (bases 1 to 556)  
AUTHORS Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.  
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE 95148729  
COMMENT On Apr 14, 1993 this sequence version replaced gi:693023.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRC, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@ldm.cl.msu.edu  
Seq primer: T7 dye primer.

FEATURES  
source

1..556  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="187517"  
/note="Vector: lambda Z1p-Lox; Site\_1: Sal; Site\_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda Z1p-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."  
BASE COUNT 151 a 118 c 123 g 146 t 18 others  
ORIGIN

Query Match 39.2%; Score 384.2; DB 23; Length 556;  
Best Local Similarity 92.9%; Pred. No. 4.9e-87;  
Matches 434; Conservative 0; Mismatches 27; Indels 6; Gaps 3;

QY 347 ctctctgacatcttgccctgagcttggaagcttgacattacagatctagagctc 406  
Db 1 CTTCTGGACATCTTGGCCGCTGAGCTTGGAGACTTGAACATTTACAGTATCAGAGCTC 60  
QY 407 tacaaaaaacatccaagaactatcctccgaacttggaattcgaagaatccatcc 466  
Db 61 TACAAAAACCAACATCCAGGACATATCTCCGAATCTGGAATCTGAAGATCTCATC 120  
QY 467 agcttgatctgtacacaacaatctacagagatagttccactctcttggaatactg 526  
Db 121 AGCTTGATCTGTACAAACAATCTTACAGGATATCTCCACTCTTGGGAAAAATTG 180  
QY 537 aagctctgtgctttttacaggttaatagacaacgattgacggtccaatccctagaag 586  
Db 181 AAGTCTCTGGCTTTTACGGCTTAATGACACCGATTGACGGGGCAATCCCTTACAGCA 240  
QY 587 ctacaggaatcccaagccttaagagtttgagcgtccaagcaatgattgtgtgaaca 646  
Db 241 CTCACGTCAATCCCAAGCCTTAAGTGTGTGATGTCTCAAGCAATGATTGTGTGNA 300  
QY 647 atcccaacaaggaaccttgcacattccctttacagaactttgagaacaacccgaga 706  
Db 301 ATCCCAACAAAGGACCTTTTGTCTCACATTCCTTTACAGACTTTGAGNACAACCCGAG 360  
QY 707 ttgg-agggacacgaattactcgtctgtcgaagctgaacactaactgacccctgaacaa 765  
Db 361 TTGGAGGACGCGAATTACTCGGTCTTGCAGCTNCGCAATTAATCAGACC---TGA 417  
QY 766 ctggcaaaacctgaanaatgaagaattggyggggtgacctgtlaagaaac 812

Db 418 ATTGAAAAACCTGAAA--TGAGATTGGGGGTGNCCTTTAAGGCAC 462

RESULT 3  
AA394359  
LOCUS AA394359 534 bp mRNA EST 30-OCT-1997  
DEFINITION 25942 Lambda-PRL2 Arabidopsis thaliana cDNA clone 305G117 3', mRNA sequence.  
ACCESSION AA394359  
VERSION AA394359.1 GI:2047570  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 534)  
AUTHORS Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.  
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE 95148729  
COMMENT On May 18, 1995 this sequence version replaced gi:811121.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRC, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@ldm.cl.msu.edu  
Seq primer: T7.

FEATURES  
source

1..534  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="305G117"  
/note="Vector: lambda Z1p-Lox; Site\_1: Sal; Site\_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda Z1p-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."  
BASE COUNT 151 a 114 c 125 g 126 t 18 others  
ORIGIN

Query Match 36.7%; Score 359.6; DB 33; Length 534;  
Best Local Similarity 89.3%; Pred. No. 8.1e-81;  
Matches 424; Conservative 0; Mismatches 44; Indels 7; Gaps 4;

QY 465 tcaactggatctgttacaacaacaatctacagagatagttccactctttgggaatat 524  
Db 1 TCACTTGATCTGTACAAACAATCTTACAGGATATCTCCACTTTTGGGAAAT 60  
QY 525 tgaagctctgtgctttttacaggttaatagacaacgattgacggtccaatccctaga 584  
Db 61 TGAAGCTCTGTGCTTTTACGGCTTAATGACAAACGATTGACCGGTCAATCCCTAAG 120  
QY 585 cactcagggcaatcccaagccttaagagttgtgagctccaagcaatgattgtgtgaa 644  
Db 121 CACTACGGCAATCCCAAGCCTTAAGTGTGTGACGTCTCAAGCAATGATTTGTGGAA 180



QY 645 caatcccaacaacgagacccttgcacattcccttacaagaacttggagaacaccga 704  
 |||||||  
 DB 181 CAATCCCAACAACGAGACCCTTGCACATTCCTTTACAGAACTTTGAGAACACCCGA 240  
 QY 705 gattggagagcaggaattactcgtcttgcagagcagacacactacacactgaaga 764  
 |||||||  
 DB 241 GATTGAGAGGACCGGAATTTCTGGCTTTCGACAGCTACGACTACCTGACAGTGA 300  
 QY 765 actgcgaacacccctgaataagaattggggggtg---acctgtgaagaacactca 821  
 |||||||  
 DB 301 ACTGCGAAACCTGGAATGAGNATTTGGGGGGGTGACCTTTGTAAAGNACACTTCANCA 360  
 QY 822 cttt--atcaataataccatctatt-atgtaataagatatatatat-gtagtaaaacaa 877  
 |||||||  
 DB 361 CTTTATTCATATTCACATTCATTCATTAAGNATTAATGATATGATGATGATGATGAT 420  
 QY 878 aaaaatgaagaatcgaatcgtataatcatcgtctcgtcgaattgaagaactcga 932  
 |||||||  
 DB 421 AAAAATGGGGATCCGATCGGGAANMNCANCGGGGCCCAATTGGAACTGNGGG 475

RESULT 4  
 LOCUS H36800 447 bp mRNA EST 30-DEC-1997  
 DEFINITION 14929 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179A227, mRNA  
 sequence.  
 ACCESSION H36800  
 VERSION H36800.1 GI:906299  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsi.

REFERENCE  
 AUTHORS 1 (bases 1 to 447)  
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
 McIntosh,L., Ohlrogge,J., Ralshel,N., Somerville,S., Thomashow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)

JOURNAL  
 MEDLINE  
 COMMENT  
 CONTACT: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313cn@dm.cl.msu.edu  
 Seq primer: T7 dye primer.

FEATURES  
 source  
 1. 447  
 Location/Qualifiers

/organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="179A227"  
 /clone\_1lb="Lambda-PRL2"  
 /note="Vector: lambda Z1p-lox. Site.1: Sal. Site.2: Not;  
 Lambda PRU2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRL's lambda Z1p-lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 0.1µg dt primed cDNA."

BASE COUNT 106 a 115 c 83 g 126 t 17 others  
 ORIGIN

Query Match 36.0%; Score 353.2; DB 23; Length 447;  
 Best Local Similarity 92.3%; Pred. No. 3.3e-79;  
 Matches 383; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

QY 13 ttaattgctctctcctcttcttgcagaaaatttcccttactcctcaattccttcg 72  
 |||||  
 DB 34 TAAATTTCTTCTTCCTTTGTTGTGAGAAATTTCCCTTACGCTCAATTCCTTTG 93  
 QY 73 attccctctcttaacctccgaagctcacatgagctcgcgaactatcgtggagct 132  
 |||||  
 DB 94 ATTTCCTCTTTAAACNCGAAGGTCACATGCGCTCGAAACTATGCTGGAGCT 153  
 QY 133 ctgcagctcgtttaaccttaaccttagcttgattaccctcgtcgaacacaccga 192  
 |||||  
 DB 154 CTTCGACCTTCCTTTAACCTTACCTTAGCTTTGATTCACCGNGTGAACAACTCCGA 213  
 QY 193 aggaatgctctctcagctcttcgcgagatttgacagatccagacatctcctcag 252  
 |||||  
 DB 214 AGGA-NTGCTCTTACGCTTTCGCCGAGTTTGACAGATCCAGACATGCTCTCCAGAG 272  
 QY 253 ctggatcccaactcttcttaactcttactcgttccatgctcctgttaacaaagaa 312  
 |||||  
 DB 273 CTGGATCCAACTCTNGTTATTCCTGTACCGGTTCCATGTANCTGTAAACCAAGCAA 332  
 QY 313 ccgctcactcgttgagattgggaattcaaacctctcgtgacatctgc-gcctgagc 371  
 |||||  
 DB 333 CCGNCTACTCTGTTGGGATTTGGGGAATTNAACCTCTTNGAGATCTTGGGGCTTAGG 392  
 QY 372 ttgggaagcttgaacatttaccagatcagaagctctcaaaaacacatccaa 426  
 |||||  
 DB 393 TTGGGAGNTGTAACATTTTCAGTATTTAGAGNTTCNCAAAACACATCCAGG 447

RESULT 5  
 LOCUS AA712221 401 bp mRNA EST 24-DEC-1997  
 DEFINITION 31449 Lambda-PRU2 Arabidopsis thaliana cDNA clone 180C97, mRNA  
 sequence.  
 ACCESSION AA712221  
 VERSION AA712221.1 GI:2722138  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsi.

REFERENCE  
 AUTHORS 1 (bases 1 to 401)  
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
 McIntosh,L., Ohlrogge,J., Ralshel,N., Somerville,S., Thomashow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)

JOURNAL  
 MEDLINE  
 COMMENT  
 CONTACT: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313cn@dm.cl.msu.edu  
 Seq primer: T7 dye primer.

FEATURES  
 source  
 1. 401  
 Location/Qualifiers

/organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="180C97"  
 /clone\_1lb="Lambda-PRU2"  
 /note="Vector: lambda Z1p-lox. Site.1: Sal; Site.2: Not;



Lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA.

BASE COUNT 103 a 105 c 78 g 99 t 16 others  
ORIGIN

Query Match 35.0%; Score 343.6; DB 37; Length 401;  
Best Local Similarity 92.1%; Pred. No. 8.8e-77;  
Matches 363; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 83 cttaacctccgaagctcacatgagctctgaactatcgtggagctcttcgagct 142  
|||||  
DB 1 ctttaaccnccgaagctcacatgagctctgaactatcgtggagctcttcgagct 60  
QY 143 tcttaaccctcaagcttagcttagcttagcttagcttagcttagcttagct 202  
|||||  
DB 61 tcttaaccctcaagcttagcttagcttagcttagcttagcttagcttagct 120  
QY 203 ctctacgctcttcgagcttagcttagcttagcttagcttagcttagcttagct 262  
|||||  
DB 121 ctctacgctcttcgagcttagcttagcttagcttagcttagcttagcttagct 180  
QY 263 actcttgtaactccttgtaactccttgtaactccttgtaactccttgtaact 322  
|||||  
DB 181 actcttgtaactccttgtaactccttgtaactccttgtaactccttgtaact 240  
QY 323 cgttgagattgggaattcaaacctctcggagattcgtcgtgagcttggaagctt 382  
|||||  
DB 241 cgttgagattgggaattcaaacctctcggagattcgtcgtgagcttggaagctt 300  
QY 383 gaacattcaagatctcgaagctctcgaagctctcgaagctctcgaagctctcga 442  
|||||  
DB 301 gaacattcaagatctcgaagctctcgaagctctcgaagctctcgaagctctcga 360  
QY 443 ctgggaattgggaattcaaacctctcggagattcgtcgtgagcttggaagctt 476  
|||||  
DB 361 ctgggaattgggaattcaaacctctcggagattcgtcgtgagcttggaagctt 393

RESULT 6  
H37300 353 bp mRNA EST 30-DEC-1997  
LOCUS 15429 Lambda-PRU2 Arabidopsis thaliana cDNA clone 17901971, mRNA  
DEFINITION  
SEQUENCE  
ACCESSION H37300  
VERSION H37300.1 GI:906799  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Arabidopsis thaliana  
thale cress.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsids.  
1 (bases 1 to 353)  
NEWMAN,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)  
95148729  
On May 8, 1995 this sequence version replaced gi:801146.  
CONTACT: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.

Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313cn@dm.cl.msu.edu  
Seq primer: T7 dye primer.  
Location/Qualifiers  
1. 353

FEATURES  
source  
/organism="Arabidopsis thaliana"  
/strain="var columbica"  
/db\_xref="taxon:3702"  
/clone="1791971"  
/note="Vector: lambda Zip-lox; Site 1: Sal; Site 2: Not;  
lambda PRU2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques). The vector is BRL's lambda Zip-lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA."  
BASE COUNT 98 a 85 c 67 g 89 t 14 others  
ORIGIN

Query Match 34.2%; Score 335.8; DB 23; Length 353;  
Best Local Similarity 95.5%; Pred. No. 8.1e-75;  
Matches 337; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 407 tacaataaacatccaaagaaactatactcttcgaaacttggaaatctgaagaatctcatc 466  
|||||  
DB 1 TACAAAACAAATCCAAAGAACTATACCTTCGAACTGGAAATCGAAGATCTCATC 60  
QY 467 agcttgatcgtgtcaacaacaatctacaggaatagttccactctcttggaaatg 526  
|||||  
DB 61 agcttgatcgtgtcaacaacaatctacaggaatagttccactctcttggaaatg 120  
QY 527 aagctctgctcttcttaacgcttaatgaacacgattgacggtccaaatccctagaga 586  
|||||  
DB 121 aagctctgctcttcttaacgcttaatgaacacgattgacggtccaaatccctagaga 180  
QY 587 ctcaaggcaatcccaagccttaagttgtgacgtctcaagcaatgattgtgtgaga 646  
|||||  
DB 181 ctcaaggcaatcccaagccttaagttgtgacgtctcaagcaatgattgtgtgaga 240  
QY 647 atcccaacaacgagcccttgctcacatctcttcaagaacttgaagaacccagaga 706  
|||||  
DB 241 atcccaacaacgagcccttgctcacatctcttgaagaacttgaagaacccagaga 300  
QY 707 ttggaaggacccggaattactcgtcttgcaagctacgacactactgacacgtg 759  
|||||  
DB 301 ttggaaggacccggaattactcgtcttgcaagctacgacactactgacacgtg 353

RESULT 7  
A1100683 364 bp mRNA EST 21-AUG-1998  
LOCUS 33804 Lambda-PRU2 Arabidopsis thaliana cDNA clone 1875XP 3', mRNA  
DEFINITION  
SEQUENCE  
ACCESSION A1100683  
VERSION A1100683.1 GI:3449393  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Arabidopsis thaliana  
thale cress.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsids.  
1 (bases 1 to 364)  
NEWMAN,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.



TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE 95148729  
COMMENT On May 5, 1995 this sequence version replaced g1:797711.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PR1, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@bm.cl.msu.edu  
The sequence entry for this EST has been reverse complemented and is being submitted in the sense orientation.  
Seq primer: M13-21.

FEATURES  
SOURCE

1. 364  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/strain="var Columbia"  
/db\_xref="taxon:3702"  
/clone="187H5XP"  
/clone\_1lb="Lambda-PR12"  
/note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not; lambda PR12 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."  
Oligo dT primed cDNA. 99 t 2 others

BASE COUNT 113 a 80 c 70 g 99 t  
ORIGIN

## Query Match

Best Local Similarity 96.4%; Score 333; DB 42; Length 364;  
Matches 350; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 506 cccactctcttggaataatgaagctctgctcttttcaaggttaataacacagattg 565  
DB 1 cccattctcttggaataatgaagctctt -gtctttttacgcttaataacacagattg 59  
QY 566 accggtccaatccctgaagcactcaagccttaacagcttgaagctcca 625  
DB 60 acggggcccaatcmtatgacacactcactgcacccaagccttaagttgattgtctca 119  
QY 626 agcaatgattctgtggaacaatcccaacaacgagccctgtccacatcccttacag 685  
DB 120 acgaaatgatttctgtggaacaatcccaacaacgagccttggctcacttctttacag 179  
QY 686 aactttgagaacaacccgagattggaaggacccggaattactcgtctctgcaagctagac 745  
DB 180 aactttgagaacaacccgagattggaaggacccggaattactcgtctctgcaagctagac 239  
QY 746 aactaattgacctggaacacacgcaaacctgaaatgaataattgggggtgacctg 805  
DB 240 actaattgacctggaacacacgcaaacctgaaatgaataattgggggtgacctg 299  
QY 806 taagaacacttcacacattatcaaatatcacatctatctgtaataagtatataatg 865  
DB 300 taagaacacttcacacattatcaaatatcacatctatctgtaataagtatataatg 359  
QY 866 agt 868  
DB 360 agt 362

RESULT 8  
Z34187/c Z34187 336 bp mRNA EST 14-JUN-1994  
LOCUS

DEFINITION ATTS3221 Versailles-VB Arabidopsis thaliana cDNA clone VBVDH02.  
MRNA Sequence.  
ACCESSION Z34187  
VERSION Z34187.1 GI:498544  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 336)  
CNRS.  
AUTHORS The Arabidopsis thaliana transcribed genome: the GDR cDNA program  
TITLE Unpublished (1996)  
JOURNAL Contact: Desprez T., Anselme J., Chiapello H., Rouze P., Caboche M., Hoite H.  
COMMENT INRA Versailles  
Laboratoire de Biologie Cellulaire  
Route de Saint-Cyr 78026 Versailles Cedex, France  
Email: thierry@versailles.inra.fr.

FEATURES  
SOURCE

1. 336  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/strain="ecotype Columbia"  
/db\_xref="taxon:3702"  
/clone="VBVDH02"  
/clone\_1lb="Versailles-VB"  
/tissue\_type="whole seedlings"  
/dev\_stage="in vitro-grown etiolated seedlings, 5 days old"  
/note="Vector: pBluescript"  
BASE COUNT 89 a 62 c 65 g 119 t 1 others  
ORIGIN

## Query Match

Best Local Similarity 99.1%; Score 328.8; DB 20; Length 336;  
Matches 330; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 633 atttggtggaacaatcccaacaacgagcccttctcacattccttaacgaacttg 692  
DB 333 atttggtggaacaatcccaacaacgagcccttctcacattccttaacgaacttg 274  
QY 693 agaacacccgagattggaaggacccggaattactcgtctctcaagctacgaactaact 752  
DB 273 agaacacccgagattggaaggacccggaattactcgtctctcaagctacgaactaact 214  
QY 753 gcaactgaaacaactggcaaacctgaataatggaattgggggtgacctgtgaagac 812  
DB 213 gcaactgaaacaactggcaaacctgaataatggaattgggggtgacctgtgaagac 154  
QY 813 acttcacacatttataataatcacatctattatgtaataagtatatatgtaagaa 872  
DB 153 acttcacacatttataataatcacatctattatgtaataagtatatatgtaagaa 94  
QY 873 acaaaaaaataagaatcgaatcggtaataatcattggtctcaattggaactcgaag 932  
DB 93 acaaaaaaataagaatcgaatcggtaataatcattggtctcaattggaactcgaag 34  
QY 933 tctgtatgtaaaattcctaattcgattcgc 965  
DB 33 tctgtatgtaaaattcctaattcgattcgc 1

RESULT 9  
A1100682 A1100682 348 bp mRNA EST 21-AUG-1998  
LOCUS A1100682  
DEFINITION 33803 Lambda-PR12 Arabidopsis thaliana cDNA clone 185L1XP 3', mRNA  
sequence.  
ACCESSION A1100682  
VERSION A1100682.1 GI:3449392  
KEYWORDS EST.  
SOURCE thale cress.



ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;  
Arabidopsids.  
1 (bases 1 to 348)  
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
Mcintosh,L., Ohlrogge,J., Ralshel,N., Somerville,S., Thomasow,M.,  
Retzel,E. and Somerville,C.  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)  
95148729  
On May 5, 1995 this sequence version replaced gi:797667.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@ldm.cl.msu.edu  
The sequence entry for this EST has been reverse complemented and  
is being submitted in the sense orientation.  
Seq primer: M13-21.  
Location/Qualifiers  
1. 348  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="185L1XP"  
/clone\_1lb="lambda-PRU2"  
/note="Vector: lambda zip-lox; Site\_1: Sal; Site\_2: Not;  
lambda PRU2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques. The vector is BRU's lambda zip-lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dt primed cDNA."  
BASE COUNT 110 a 74 c 66 g 93 t 5 others  
ORIGIN  
Query Match 32.2%; Score 315.6; DB 42; Length 348;  
Best Local Similarity 96.4%; Pred. No. 1e-69;  
Matches 321; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

RESULT 10  
AM443205 608 bp mRNA EST 17-FEB-2000  
LOCUS  
DEFINITION  
EST308135 tomato mixed elicitor, BRT Lycopersicon esculentum cDNA  
clone CLEF43M10 5', mRNA sequence.  
ACCESSION  
AM443205  
VERSION  
AM443205.1 GI:6983387  
KEYWORDS  
EST.  
SOURCE  
tomato.  
ORGANISM  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
1 (bases 1 to 608)  
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,  
Liang,F., Hansen,T.S., Rohnig,C.M., Craven,M.B., Bowman,C.L.,  
Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,  
Giovannoni,J.J. and Martin,G.B.  
Generation of ESTs from tomato callus (mixed elicitor)  
Unpublished (1999)  
On Dec 20, 1995 this sequence version replaced gi:1135853.  
Contact: David Frisch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4366  
Fax: 864 656 4293  
Email: dfrisch@clemson.edu  
5 prime sequence.  
Location/Qualifiers  
1. 608  
/organism="Lycopersicon esculentum"  
/cultivar="Rio Grande Fluor"  
/db\_xref="taxon:4081"  
/clone="CLEF43M10"  
/clone\_1lb="tomato mixed elicitor, BRT"  
/tissue.type="leaf"  
/dev stage="4-6 week old plants"  
/lab host="X11-Blue MR"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; cLEF - inoculated with a variety of disease response  
elicitors. Plants exposed to 2,6 dichloroisonicotinic  
acid, BTH, jasmonic acid, ethylene, fenthion, Etx,  
okadaic acid, or systemin prior to tissue harvest. EcoRI  
site was destroyed during cloning."  
BASE COUNT 165 a 126 c 134 g 183 t  
ORIGIN  
Query Match 31.9%; Score 312.6; DB 81; Length 608;  
Best Local Similarity 71.5%; Pred. No. 6.2e-69;  
Matches 411; Conservative 0; Mismatches 164; Indels 0; Gaps 0;



Db	273	CTTAAACATCTACAGTATCTGAGACTTTACAAAAATAATATCACGGAAACATCCCTAAG	332
QY	440	gaacttggaaatctgaagaatctcatcagcttggatctgtacaacaacaatcttaagg	499
Db	333	GAGCTCGCTAACTTGAAGAGCGTTATTAAGTCTGGATCTGTACAAACAACAATATTTGGGG	392
QY	500	atagttccactcttcttgggaaattgaagtcctctgtctctttaaagcttaatgaac	559
Db	393	ACAATTCCTACTTACTTCTGGAAACCTTAAAAACCTTTTCTTTTGCGCTTAAATGATAAC	452
QY	560	cgattgaccggtcccaatcccttagacgcactacggycaatcccaagccttaagtgtgac	619
Db	453	ANGTTAACAGGACCAATGCCAAGAAGAACTTACTAGCATTTCTAGCCGTGAAGAATGTGGAT	512
QY	620	gtctcaagcaatgtattgtgtgtgaacaatcccaacaagaagacccttgcctacatcct	679
Db	513	GTCCTCGATATGATTTGTGTGTGGACAACTTCCTCTCTGGTCCATTGTGAGCATATTCCT	572
QY	680	ttacagaactttgagacaacccgagatttggagg	714
Db	573	CTAAACATTTTCGAGCACAAATCCTCGACTTGAAG	607

RESULT	11
LOCUS	AM031110
DEFINITION	AM031110 720 bp mRNA EST 15-SEP-1999 EST274417 tomato calyx, TMU Lycopersicon esculentum cDNA clone cLEC13C01 similar to leucine-rich repeat protein (LRR), mRNA sequence.
ACCESSION	AM031110
VERSION	AM031110.1
KEYWORDS	GI:5889866
	EST.

ORGANISM  
Lycopersicon esculentum  
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core  
eudicots: Asterales; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.

REFERENCE	TITLE	JOURNAL	COMMENT
1 (Pages 1 to 720) Alcala, J., Vrehalov, J., White, R., Matern, A.L., Viston, T., Holt, I.E., Liang, F., Upson, J., Craven, M.B., Bowman, C.L., A Romning, C.M., Fraser, C.M., Martin, G.B., Tanksey, S.D., and Giovannoni, J.	Generation of ESTs from tomato callus tissue	Unpublished (1999)	On Dec 20, 1999 this sequence version replaced g1:1133546.

**TITLE** Generation of ESTs from tomato callus tissue  
**JOURNAL** Unpublished (1999)  
**COMMENT** On Dec 20, 1995 this sequence version replaced  
Contact: David Ertisch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel.: 864 656 4366  
Fax: 864 656 4293  
Email: dirtsch@CLEMSON.EDU  
5 prime sequence.

FEATURES	Location/Qualifiers
source	1. .720

```

/organism="Lycopersicon esculentum"
/cultivar="LA96"
/db_xref="taxon:4081"
/clonename="CLB13C21"
/clonename_11b="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="Xll Blue MRF/n"
/notes="Vector: pBluescript SK(-): 5

```

BASE COUNT	200 a	149 c	143 g	228 t
ORIGIN				

Query Match	31.8%;	Score 311.6;	DB 64;	Length 720;
Best Local Similarity	71.4%;	Pred. No. 1.1e-68;		
Matches 410; Conservative	0;	Mismatches 164;	Indels 0;	Gaps 0

Qy	140	gcttcgttaacccttaaccttaagctttagattcaacctgctgcgaaganaaacctccgaagagat	139
Qy	141		
Db	147	gatttttgscagattgtcttctgtacotgngcctgtaaaagggaatttcagaaggsggat	206
Qy	200	gctcttaagcctcttcgcgcggaagttagacagatccagaccatgctccctccagatgggat	259
Qy	201		
Db	207	gcatttgacccctccgcgcggagcttatctgcacccggctaacgctttacagagctgggat	266
Qy	260	ccaaacttbttaacctctgtacacctggtctccatgycacctgtaaccaagaacccgctc	319
Db	267	ccaaattctgttaaaccttctgtaacctggtttcatgtcactttgcamaoggaaganaatcagagt	326
Qy	320	actcgttggatttgggaaattcaaacctctctgtgacacttgcgcctbagcttgggaag	379
Db	327	actcgttggatttgggaaattcaaacctctctgtgacacttgcgcctbagcttgggaag	386
Qy	380	cttgaacatttaagatattagatctagagctctacaaaacaacatccagaagactatacctcc	439
Db	387	cttgaacatttaagatattagatctagagctctacaaaacaacatccagaagactatacctcc	446
Qy	440	gaacttggaaatctgagaatctcatcagcttggatctgtlcaacaacaaatcttacagg	499
Db	447	gagctgggttaactggaagagccttattatgtcgtatctgtacaaacaaatattttccggg	506
Qy	500	atagttcccaactctcttgggaaatattgaagctctcgtgccttttaacggttaagaaca	559
Db	507	acaatttctcttcttcaacttggaaaacctggaataacctgtttcttgcgttaaatgatatac	566
Qy	560	cgattgaccggtccaaatcccttagagcaactcacaggaatcccaagccttaaaagtgttgc	619
Db	567	aagcttaacagagaccatcccaaaagaaacttactagactttttagccttaaaagttgtgat	626
Qy	620	gtctcaagcaatgatttgttggagaacatcccaacaaacgagacccttgcctcaatctc	679
Db	627	gtctcgatattatgtttgtgtgagaaataattcttacttctgtgcattttgacacataattct	686
Qy	680	ttacagaacttggagaacacccgagatgggg	713
Db	687	cttaaacatttcgagcacaatctctcactttgaaag	720

RESULT	12				
LOCUS	234606	337	bp	MRNA	EST
DEFINITION	AT5G3366 Versailles-VB Arabidopsis thaliana cDNA clone VBDH02 MRNA sequence.				

VERSION 234606.1 GI:507009

SOURCE      thale cress.  
ORDINARY      Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Streptophyta; Charophyceae; Lemnaceae

eudicots; Rosidae;  
Arabidopsis.

REFERENCE 1 (bases 1 to 337)  
AUTHORS CNRS.

THE  
JOURNAL  
Unpublished (1996)

M., Hofte H.

Laboratoire de Biochimie

Email: thierry@vers

source	1. .337
--------	---------

FEATURES	Location/Qualifiers
source	1. .337



/organism="Arabidopsis thaliana"  
 /strain="ecotype Columbia"  
 /db\_xref="taxon:3702"  
 /clone="YBYDH02"  
 /clone\_lib="Versailles-VB"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="in vitro-grown etiolated seedlings, 5 days old"  
 /note="vector: pBluescript"  
 /note="vector: pBluescript"

BASE COUNT 73 a 98 c 60 g 105 t 1 others

ORIGIN

Query Match 31.1%; Score 305; DB 20; Length 337;  
 Best Local Similarity 97.6%; Pred. No. 56-67;  
 Matches 330; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

17 ttgtcttccctctgttcagaaatttcccttacttcaaatcttctgatt 76  
 |||||  
 Db 1 TTGCTTCTCCCTTGTTCAGAAATTTCCCTTACTTCGAATTCCTTTGATT 60  
 |||||  
 QY 77 cccctcttaacctcgaagctcacatggcgctcgaactatcggtggagctcttc 136  
 |||||  
 Db 61 CCCCTCTTAAACCTCGAAGCTCAATGGCGCTCGAACTATCGGTGGAGCTCTTC 120  
 |||||  
 QY 137 gcaagcttgtaacccctaaccttagcttgattcaccttgctgaaagcaactcgaagga 196  
 |||||  
 Db 121 GCACCTTCGTAACCTCAACCTTAGCTTTCATTACCTGTCGAACCAACTCCANA-GA 179  
 |||||  
 QY 197 gatgctctcagctcttcgacgagatttgacagatccagacatgctctccagaagctgg 256  
 |||||  
 Db 180 ATGCTNCTCAAGCTCTTCGCCGAGATTGACAGATCCAGACCATGTCTCCAGAGCTGG 239  
 |||||  
 QY 257 gatccaaatttgtaaccttgtaaccttgtaaccttgtaaccttgtaacaaacacgcg 316  
 |||||  
 Db 240 GATCAACCTTGTGTAATCTTGTAACCTTGTAACCTTGTAACCTTGTAACCAAGCAACGC 299  
 |||||  
 QY 317 gtacatcggtg-ggattgggaattcaaacctctctg 353  
 |||||  
 Db 300 GTCACCTGCTGGGATTGGGAATTCAAACCTCTCTG 337

RESULT 13  
 LOCUS A1100679 339 bp mRNA EST 21-AUG-1998  
 DEFINITION 33800 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179j11xp 3', mRNA  
 sequence.  
 ACCESSION A1100679  
 VERSION A1100679.1 GI:3449389  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsids  
 1 (bases 1 to 339)  
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
 McInosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)

TITLE  
 JOURNAL 95146729  
 MEDLINE  
 COMMENT On May 5, 1995 this sequence version replaced g1:797664.  
 Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313c@elm.cl.msu.edu  
 The sequence entry for this EST has been reverse complemented and

is being submitted in the sense orientation.  
 Seq primer: M13-21.  
 Location/Qualifiers  
 1..339  
 source  
 /organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="YBYDH02"  
 /clone\_lib="Versailles-VB"  
 /clone\_lib="Lambda-PRL2"  
 /note="vector: lambda ZAP-Lox; Site\_1: Sal; Site\_2: Not;  
 Lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques). The vector is BLU's lambda ZAP-Lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dt primed cDNA."

BASE COUNT 104 a 75 c 66 g 89 t 5 others

ORIGIN

Query Match 30.4%; Score 298.2; DB 42; Length 339;  
 Best Local Similarity 94.6%; Pred. No. 2,66-65;  
 Matches 317; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 534 tggctctttaaaggcttaagtgacaaacgattgacgcgacaa-tccctagagcaactca 592  
 |||||  
 Db 2 TGTCTTTTACGGCTTAATGACAAACCGATTGACGGGGCAATTCCTTAGAGCACTACT 61  
 |||||  
 QY 593 gcaaccacaagccttaagttgttgacgtctcaagcaatgtttgtggaacatccca 652  
 |||||  
 Db 62 GCATTCCAAAGCTTAAATTTGTGATNTCTCAACAATGATTTGTGGAACANTCCA 121  
 |||||  
 QY 653 acaacggagcccttgctcacatctccttcaagaaatttgagaaacccgagattggag 712  
 |||||  
 Db 122 ACAACGGAGCACTTTGCTCACATTCCTTACAGAACTTTGAGAACACCCAGGTTGGAG 181  
 |||||  
 QY 713 ggaacggaattactggttcttgaagctgacgacactaactgacactgaaacactggaa 772  
 |||||  
 Db 182 GGAACGGGAATTACTCGGCTTGCAAGCTGACACTACTGCACTGAAATAATTTGGCA 241  
 |||||  
 QY 773 aactgaaatgaagaattggggggtgacctgtgtaagaacacttaccacttacaat 832  
 |||||  
 Db 242 AANCTGAATAATGAAGATTGGGGGGTGACCTGTGAAGAACCTTACCACTTTGCAAAAT 301  
 |||||  
 QY 833 atcaatctattatgtaataagatatatgtag 867  
 |||||  
 Db 302 ANCAATCTACTATGTAAATGAATATATATGTAG 336

RESULT 14  
 LOCUS AM221939 646 bp mRNA EST 07-DEC-1999  
 DEFINITION EST298750 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA  
 clone cENEM3, mRNA sequence.  
 ACCESSION AM221939  
 VERSION AM221939.1 GI:6533623  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 646)

REFERENCE  
 AUTHORS Alcala,J., Vrabalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,  
 Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,  
 Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and  
 Giovannoni,J.  
 Generation of ESTs from tomato fruit tissue  
 JOURNAL Unpublished (1999)



COMMENT On Jul 8, 1999 this sequence version replaced gi:5422558.

Contact: David Ertisch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4366  
Fax: 864 656 4293  
Email: dertisch@CLEMSON.EDU  
5 prime sequence.

FEATURES  
Location/Qualifiers  
1. 646  
source

/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone\_1lb="tomato fruit red ripe, TAMU"  
/issue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopen accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe), 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

BASE COUNT 181 a 126 c 138 g 201 t

ORIGIN

Query Match 29.8%; Score 293.8; DB 74; Length 646;

Best Local Similarity 71.7%; Pred. No. 3.6e-64;  
Matches 385; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 140 gcttcgttaaccccttaacctagcttgcagcgaagcaactccgaaggagat 199  
DB 110 GGTGTGGCAGTGTCTTGTGCTGTAGCTGTGGCTGTAAAGGGAATCAGAAGGGAT 169  
QY 200 gctctcaagctcttcgcggagatttcacagatccagacatctcccgagctggat 259  
DB 170 GCTTTGACGCCCTCCCGAGAGCTTATCTGACCCGGGTAACGTTTACAGAGCTGGGAT 229  
QY 260 ccaactctgttaacctctgtacgtgttcacgtacccgttaaccagaacacccgcctc 319  
DB 230 CCAATCTGTGTAACCTTGTACTGTCTGATCTGACGTAACGAGATTAATCAAGT 289  
QY 320 actcgtgtaatttggaaattcaaacctctcgtgacatcttgcgcctgagcttggagag 379  
DB 290 ACTGCTGTGATCTTGGGAAGTCAAAAGTTATCTGTCATTTGGTATCTGAGCTCGGAAAG 349  
QY 380 cttaagaatttagatcttagagctctcaaaaaaacatccagaagactatactcc 439  
DB 350 CTTGAACCTTCAAGTATCTGGAGCTTACAAAATTAATTCAGGGAACCATCCCTAAG 409  
QY 440 gaacttggaaattcgaagaatcctacagcttgcgtacacaaacaactcttacaggg 499  
DB 410 GAGCTCGGTACTTGAAGAGCCTTATAGTGTGATCTGTACACAAATATATTCGGGG 469  
QY 500 atagctccactctcttgggaaattgaagctctgtgcttttaaggcttaatgacac 559  
DB 470 ACAATCTCTACTTACTTGGAAACCTGAAAACCTTGTTTCTGCTCTAAGAAAGATAC 529  
QY 560 cgattgacggttcaatccctgaagcaactcaagcaatcccaagcttaagttgttgc 619  
DB 530 AAGCTAACAGACCAATCCCAAGAGACTTACTAGCATTTTACTCTGAAAAGTTGTGAT 589  
QY 620 gtctcaagcaattgttgcgtgaaacaatcccaacaacggaaccttgcacatt 676  
DB 590 GTCGTGAATTAATGATTGTGTGAACAATTCCTACTCTGTGATTTGACATATT 646

RESULT 15

AM350720/c AM350720 657 bp mRNA EST 01-FEB-2000

LOCUS GM210009A10P4R Gm-r1021 Glycine max cDNA 3', mRNA sequence.

DEFINITION AM350720  
VERSION AM350720.1 GI:6848430

KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
Glycine.

1 (bases 1 to 657)

Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Rhanna, A., Coryell, V.,  
Expelling, J., Raph, C., Shoop, E., Pardinas, J., Liu, L., and Lewin, H.,  
A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)  
On Nov 22, 1999 this sequence version replaced gi:5462107.

Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics

University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147  
Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu  
This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134. For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or info@genome  
systems.com web site: www.genomesystems.com

Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.  
Location/Qualifiers

FEATURES  
Source

1. 657  
location/Qualifiers

/organism="Glycine max"  
/cultivar="Williams"

/db\_xref="taxon:3847"  
/clone\_1lb="Gm-r1021"

/issue\_type="root"  
/lab\_host="XL10-Gold"

/note="Vector: pBluescript II XR. Site.1: EcoRI; Site.2:  
XhoI; Library Gm-r1021 is a sequence-driven, resequenced set  
of the original library Gm-cl04 which was prepared from  
root cDNA. The mRNA was isolated from entire roots of 8  
day old 'Williams' seedlings which were propagated on  
paper towels with distilled water. StrataGene's cDNA  
Synthesis Kit (catalog #200401) was used to synthesize  
the cDNA. The Gm-cl04 library was constructed by Dr.  
Paul Keim & Virginia H. Coryell, Department of Biology,  
Box5640, Northern Arizona University, Flagstaff, AZ  
86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu.  
The contig analysis to select unique genes was performed  
by the laboratory of Ernest Retzel, Computational Biology  
Centers, University of Minnesota,

http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html  
Rearranging was performed by Genome Systems, St. Louis,  
http://www.genomesystems.com, and sequencing by the Keck  
Center for Comparative and Functional Genomics,  
University of Illinois,  
http://www.life.uiuc.edu/biotech/keck.html."

BASE COUNT 190 a 131 c 120 g 191 t 25 others  
ORIGIN

Query Match 29.8%; Score 292.8; DB 80; Length 657;

Best Local Similarity 72.4%; Pred. No. 6.4e-64;  
Matches 360; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 272 aacctgtactgttcatcgtacacgttaaccagaacacggtcactcgttggat 331  
DB 657 AGTCCTGTACANNNTTCCAGCTCANNNGCAANCGAGCAANNNTTGTACCTCGATGAT 598



Oy	332	ttgggaattcaaacccctctggaccattctggccctgagcttggagcttggaaacttgaacatta	391
Db	597	NNTGGTACTCTTAACCTTACTTCGGACATTTGGTACNNNGAATTGGGAAGCTGGAGCATCTA	538
Oy	392	cagtatctagagctctcaaaaaaacatccaaaggaactaactctccgaacttggaaat	451
Db	537	CAGTATCTTGAGTTGTACAAAACAAACATTCAGGAACCTATTNNNNNGAACCTTGGAAAC	478
Oy	452	ctgaagaatctcatcagcttggatctgttaacaacaacaattctcagggaatgttccaat	511
Db	477	CTGAAGACTCTAGTTACTCTTGGAATTTGTACACACACAACTATCAGGACACCACTCCACCT	418
Oy	512	tctttgggaanaattggaagctcctgctcttctttaaaggcttaatgacaacgattgaccggt	571
Db	417	TCATTGGGGGAATTGAAGAACTCTGTCTTTTACAGATCAATGACACACGACCTAAATGGC	358
Oy	572	ccaatcccttagagcactcaagcgcaatcccaagccctaaagtgtgtgaagctcgaagcaat	631
Db	357	CCAAATTCCCAAGGAAGCTGTGCGCTGTTCAAAGCCTTTAACTAGTGGATGCTCTCCAAAT	298
Oy	632	gatttggttgaanaaattcccaacaaagggcccttctcaagttctcttaaggaactt	691
Db	297	GATTTATGTGGTACAAATTCCTACATCTGGGCCATTGACAGATTTTCATTGAAATTACTTT	238
Oy	692	gagaacaaccccgagatttggaggagccggaatttactcgtctcttgcgaactgaacataac	751
Db	237	GAGATATATCCCCGCTTGGAAGTCCACAGAGTTGTTGGACTATGAAATATATGACAAAC	178
Oy	752	tgcaccttgaanaacactg 768	
Db	177	TGCTGTGAATTAACGTG 161	

Search completed: June 23, 2000, 19:06:35  
Job time: 27428 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:19:09 ; Search time 12463.9 Seconds

(without alignments)  
-61.580 Million cell updates/sec

Title: US-09-180-798-26

Sequence: 1 tccgaccacgcgtccgcgaa.....agtccaaaaaaaaaaaaa 789

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_da3:\*  
4: gb\_da4:\*  
5: gb\_da5:\*  
6: gb\_da6:\*  
7: gb\_da7:\*  
8: gb\_da8:\*  
9: gb\_da9:\*  
10: gb\_da10:\*  
11: gb\_da11:\*  
12: gb\_da12:\*  
13: gb\_da13:\*  
14: gb\_da14:\*  
15: gb\_da15:\*  
16: gb\_da16:\*  
17: gb\_da17:\*  
18: gb\_da18:\*  
19: gb\_da19:\*  
20: gb\_da20:\*  
21: gb\_da21:\*  
22: gb\_da22:\*  
23: gb\_da23:\*  
24: gb\_da24:\*  
25: gb\_da25:\*  
26: gb\_da26:\*  
27: gb\_da27:\*  
28: gb\_da28:\*  
29: gb\_da29:\*  
30: gb\_da30:\*  
31: gb\_da31:\*  
32: gb\_da32:\*  
33: gb\_da33:\*  
34: gb\_da34:\*  
35: gb\_da35:\*  
36: gb\_da36:\*  
37: gb\_da37:\*  
38: gb\_da38:\*  
39: gb\_da39:\*  
40: gb\_da40:\*  
41: gb\_da41:\*  
42: gb\_da42:\*  
43: gb\_da43:\*  
44: gb\_da44:\*

45: gb\_da45:\*  
46: gb\_da46:\*  
47: gb\_da47:\*  
48: gb\_da48:\*  
49: gb\_da49:\*  
50: gb\_da50:\*  
51: gb\_da51:\*  
52: gb\_da52:\*  
53: gb\_da53:\*  
54: gb\_da54:\*  
55: gb\_da55:\*  
56: gb\_da56:\*  
57: gb\_da57:\*  
58: gb\_da58:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	789	100.0	789	5	A67821	A67821 Sequence 26
2	769.6	97.5	1063	5	A67825	A67825 Sequence 30
3	742	94.0	981	5	A67819	A67819 Sequence 24
4	714.2	90.5	1106	5	A67817	A67817 Sequence 22
5	361	45.8	894	5	A67823	A67823 Sequence 28
6	211.8	26.8	2089	5	A67827	A67827 Sequence 32
7	155	19.6	936	8	SBU62279	U62279 Sorghum b1c
8	103.4	13.1	1755	8	DCU93048	U93048 Daucus caro
9	103.4	13.1	1815	5	A67797	A67797 Sequence 2
10	99	12.5	4604	7	LEIRPGENE	X95269 L.esculentu
11	84.8	10.7	101284	8	ATAC009991	AC009991 Arabidops
12	81.6	10.3	116944	8	AC008017	AC008017 Arabidops
13	81.4	10.3	5033	8	IN077888	U77888 Ipomoea nil
14	79.8	10.1	81513	50	AC021198	AC021198 Arabidops
15	79.2	10.0	134402	8	ATAC011765	AC011765 Arabidops
16	78.8	10.0	100887	50	ATAC011620	AC011620 Arabidops
17	78.8	10.0	108355	50	AC005957	AC005957 Arabidops
18	78.6	10.0	133840	8	AC003981	AC003981 Complete
19	76.8	9.7	112529	44	AC015446	AC015446 Arabidops
20	76.4	9.7	23143	8	F15P11	AF160760 Arabidops
21	75.8	9.6	142418	7	AP000815	AP000815 Oryza sat
22	75.4	9.6	3176	5	E12705	E12705 Arabidops
23	75.4	9.6	3176	8	ATU47029	U47029 Arabidops
24	75.4	9.6	3176	8	OLU72727	U72727 Oryza longi
25	75.4	9.5	5940	7	AB010075	AB010075 Arabidops
26	74.6	9.5	84440	7	AB010075	AB010075 Arabidops
27	72	9.1	87434	7	AB010698	AB010698 Arabidops
28	71.4	9.0	3573	5	A57133	A57133 Sequence 4
29	71.4	9.0	5368	8	U42445	U42445 Lycopersico
30	71.4	9.0	6471	5	A57130	A57130 Sequence 1
31	71.4	9.0	6471	8	U42444	U42444 Lycopersico
32	71.4	9.0	61712	7	AB017061	AB017061 Arabidops
33	70.8	9.0	4081	5	A67815	A67815 Sequence 20
34	70.8	9.0	6695	5	A67796	A67796 Sequence 1
35	70.8	9.0	98471	8	AC012654	AC012654 Arabidops
36	70.6	8.9	2278	8	SBRLK1	Y14600 Sorghum b1c
37	70.6	8.9	2786	7	AB029327	AB029327 Nicotiana
38	70.6	8.9	3979	5	A67428	A67428 Sequence 1
39	70.6	8.9	3979	5	A67429	A67429 Sequence 2
40	70.6	8.9	3979	6	AF053993	AF053993 Lycopersi
41	70.6	8.9	4123	5	A67434	A67434 Sequence 7
42	70.6	8.9	4123	8	AF053998	AF053998 Lycopersi
43	70.2	8.9	9424	8	OSU72724	U72724 Oryza sativ
44	70.2	8.9	52717	7	AB019227	AB019227 Arabidops
45	70	8.9	101284	8	ATAC009991	AC009991 Arabidops

ALIGNMENTS



```

RESULT 1
LOCUS A67821 789 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 26 from Patent WO9743427.
ACCESSION A67821
VERSION A67821.1 GI:4756643
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 789)
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source Location/Qualifiers
1..789
/organism="unidentified"
/db_xref="taxon:32644"
<2..664
/feature="unidentified"
/codon_start=1
/protein_id="CAB42248.1"
/db_xref="GI:4756644"
/translation="RRPRPNRYRWELEFASLITITLALIHVEANSRSDALYALRSIT
DPDHLQSDPFLVNPCTFHYTCNDNRYTRVDLNSNLSGLAPLELKEHLYLE
LYKNNIGTIPSELGNLKLISLDLNNNLGIVPISLGLKSLVFLRLNDRNLTEPI
PRAITAIPLSLKVVDSNDLCGTFPTNGPFAHPIPLQNEENRPLRGLGLASYDIN
CT"
BASE COUNT 235 a 191 c 156 g 207 t
ORIGIN
Query Match 100.0%; Score 789; DB 5; Length 789;
Best Local Similarity 100.0%; Pred. No. 1.9e-215;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tcgaccacgagtcgagaaactatcgtagagctcttcgacgttcgttaacccaaac 60
DB 1 tgcagccacgagtcgagaaactatcgtagagctcttcgacgttcgttaacccaaac 60
QY 61 cttaagcttgaattcaccgctgcgaagcaaacctcgaagagaatgctcttaccgcttcg 120
DB 61 cttaagcttgaattcaccgctgcgaagcaaacctcgaagagaatgctcttaccgcttcg 120
QY 121 ccgagagtttaagaagctccggaacatgcttcgaagctggagatccaaactctgttaacc 180
DB 121 ccgagagtttaagaagctccggaacatgcttcgaagctggagatccaaactctgttaacc 180
QY 181 ttgtaccgtgttcacatgacacgtgaacaaagacacacgctgaactcgtgtgattggg 240
DB 181 ttgtaccgtgttcacatgacacgtgaacaaagacacacgctgaactcgtgtgattggg 240
QY 241 gaattcaaacctcctcgtgacatcttgccctgagcttgggaagcttgaacattcaagta 300
DB 241 gaattcaaacctcctcgtgacatcttgccctgagcttgggaagcttgaacattcaagta 300
QY 301 tcttaagctctcaaaaacacacatccaaggaactatccttcgaagcttgaagaaatcgaa 360
DB 301 tcttaagctctcaaaaacacacatccaaggaactatccttcgaagcttgaagaaatcgaa 360
QY 361 gaattcaccagcttgaattctgttaacaacacatcttaacaggaatgcttcccaattcctt 420
DB 361 gaattcaccagcttgaattctgtgttaacaacacatcttaacaggaatgcttcccaattcctt 420
QY 421 gggaaaaatgaaagtctcgtgtcttttaagcgttaatacaacacgattgaaggggccaat 480
DB 421 gggaaaaatgaaagtctcgtgtcttttaagcgttaatacaacacgattgaaggggccaat 480
QY 481 ccctagagcaactcactgcaatcccaagccttaagcttgaatgcttgaatgaatgatt 540
DB 481 ccctagagcaactcactgcaatcccaagccttaagcttgaatgcttgaatgaatgatt 540

```

```

QY 541 gtytgaaacaatcccaacaacaggaaccttgcgtcacatcccttaccagaacttgaaga 600
DB 541 gtytgaaacaatcccaacaacaggaaccttgcgtcacatcccttaccagaacttgaaga 600
QY 601 caaccagagcttgaagaggaacgaattactcgtcgtcgaagcgaacgaactgaacac 660
DB 601 caaccagagcttgaagaggaacgaattactcgtcgtcgaagcgaacgaactgaacac 660
QY 661 ctgaaaaaatcggcaaaacccctgaagaatctgggggggctgaacacttgaagaacatt 720
DB 661 ctgaaaaaatcggcaaaacccctgaagaatctgggggggctgaacacttgaagaacatt 720
QY 721 caacccttaccatcatcatctactatgtaataatataatataatgtaacataaaaaa 780
DB 721 caacccttaccatcatcatctactatgtaataatataatataatgtaacataaaaaa 780
QY 781 aaaaaaaaa 789
DB 781 aaaaaaaaa 789
RESULT 2
LOCUS A67825 1063 bp DNA PAT 05-MAY-1999
DEFINITION Sequence 30 from Patent WO9743427.
ACCESSION A67825
VERSION A67825.1 GI:4756647
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1063)
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.
TITLE PRODUCTION OF APOMITIC SEED
JOURNAL Patent: WO 9743427-A 20-NOV-1997;
CIBA GEIGY AG (CH)
FEATURES
source Location/Qualifiers
1..1063
/organism="unidentified"
/db_xref="taxon:32644"
106..762
/feature="unidentified"
/codon_start=1
/protein_id="CAB42252.1"
/db_xref="GI:4756648"
/translation="MASRNRWELFASLITITLALIHVEANSRSDALYALRSITDP
DPDHLQSDPFLVNPCTFHYTCNDNRYTRVDLNSNLSGLAPLELKEHLYLE
LYKNNIGTIPSELGNLKLISLDLNNNLGIVPISLGLKSLVFLRLNDRNLTEPI
PRAITAIPLSLKVVDSNDLCGTFPTNGPFAHPIPLQNEENRPLRGLGLASYDINCT"
BASE COUNT 313 a 242 c 206 g 302 t
ORIGIN
Query Match 97.5%; Score 769.6; DB 5; Length 1063;
Best Local Similarity 98.9%; Pred. No. 6.9e-210;
Matches 775; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 5 ccaacgctccgagaaactatcgtagagctcttcgacgttcgttaacccaaac 64
DB 103 ccaacgctccgagaaactatcgtagagctcttcgacgttcgttaacccaaac 162
QY 65 gcttaccacgtgctgaagcaaacctcgaagagaatgctcttaccgcttcgcgg 124
DB 163 gcttaccacgtgctgaagcaaacctcgaagagaatgctcttaccgcttcgcgg 222
QY 125 agttaacagatccggaacatgcttcgaagcgtggagatccaaactctgttaacacctgt 184
DB 223 agttaacagatccggaacatgcttcgaagcgtggagatccaaactctgttaacacctgt 282
QY 185 acctggtcactgtaaccctgtaacaaagacacacgctgaactcgtgtgattgggaat 244

```



Db 283 ACCTGTTCCATGTCACCTGTAACCAAGACACCGGCTACCTGCTGTGATTTGGGGANT 342  
Qy 245 tcaaacctctctgacacatcttgcgcctgagcttggaaagcttgaacattacatcatc 304  
Db 343 TCAAACTCTCTGGACATCTTGGCCGAGACTGGGAAGCTTGAACATTATACAGATATCA 402  
Qy 305 gagcttcaaaaaaacaacatccaaagaaactacatcccttcgaaacttggaaatcgaagaat 364  
Db 403 GAGCTTCAAAAAACAACATCCAAAGAACTATCTTCGAACTTGGAAATCTGAAGAAAT 462  
Qy 365 ctcatcagcttgatctgtcaacaacaacatcttcaagaggaatgtccatctcttggga 424  
Db 463 CTGATCAGCTTGATCTGTACAAACAATCTTACAGGAGTAGTCCACTCTTTGGGA 522  
Qy 425 aaattgaagctctgtgtcttcttaaggcttaataagacaacggatgagcggggccaatccct 484  
Db 523 AAATTGAAGCTCTGTGCTTTTACCGCTTAAAGACACGATGAGGGGCAATCCCT 582  
Qy 485 agagcactcaatgcaatcccaagccttaaaagtgttgatgtcgaagcaatgattgtgt 544  
Db 583 AGACACTCACTGCAATCCCAAGCCTTAAAGTTGTGATGTCTCAAGCAATGATTTGTGT 642  
Qy 545 ggaacaatcccaacaagaagaccttctgcacatcccttcaagaacttgaagaacac 604  
Db 643 GGAACAATCCCAACAAGACCTTTTGTCTACATTCCTTTACAGAACTTTGAGAAACAAC 702  
Qy 605 ccgaggttggagggagccggaattactcgtgtcttgcgaagctgaagcaactacgtcacctga 664  
Db 703 CCGAGGTTGAGGAGCCGGAATTAAGTGTCTTGCAGAGCTACACACTTACCTGACCTGA 762  
Qy 665 aaaaattggcaaaaacctgaaatgaagaatttgggggttgaaccttgaagaacacttcac 724  
Db 763 AAAAATTGGCAAAACCTGAAATGAGAAATGGGGGTGACCTTGTAAACAACCTTCACAC 822  
Qy 725 accttacaatatcacatctactatgttaataagtatataatgtagtccaaaaaataa 784  
Db 823 ACTTATCAAAATATCACATCTACTATATTAATGATATATGATGATCCAAAAAATAA 882  
Qy 785 aaaa 788  
Db 883 TGA 886

RESULT 3  
LOCUS A67819 981 bp DNA PAT 05-MAY-1999  
DEFINITION Sequence 24 from Patent WO9743427.  
ACCESSION A67819  
VERSION A67819.1 GI:4756641  
KEYWORDS  
SOURCE .  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 981)  
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE PRODUCTION OF APOMICRITIC SEED  
JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
source location/Qualifiers  
1..981  
/organism="unidentified"  
/db\_xref="taxon:32644"  
104..760  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CA842246.1"  
/db\_xref="GI:4756642"  
/translation="MASRNTWELFPAASITLTLALIHVEANSESDALYALRSITDP  
DHVQSDPTLVNPTWTFHYTCQNDNRVTRVDLGSNSLGHAPLGLKLEHLYTELY  
KNNIQGIIPSELGNLNLISLDYNNNLGIPTSLGKLSKSLVFLRNDNRITGPIPR  
ALTAIPSLKVVYDVSSNDLCGTIPTNGPFAHIPIDLOFNENPRLEGPGLIASYDNTCT

BASE COUNT 286 a 236 c 180 g 279 t

ORIGIN  
Query Match 94.0%; Score 742; DB 5; Length 981;  
Best Local Similarity 96.8%; Pred. NO. 5.5e-202;  
Matches 757; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
Qy 5 cccacgcgtccggaacatcgttggagagctctgcagcttcttgaatccttaacctta 64  
Db 101 CACATGGCGCTCTGAAACATGCGTGGAGCTCTCCGACGCTTGTTAACCTTAACCTTA 160  
Qy 65 gctttatcaccgttgcgaagcaaacctccgaagagatgtctcttcagccttcgcgcg 124  
Db 161 GCTTGAATTCAGTGTGTCAGAGAAACTCCGAGGAGATGCTCTCAGGCTCTTGGCCGG 220  
Qy 125 agttacagatccggaacatgtcttcagagcttggatcccaactctgttaaccttgt 184  
Db 221 AGTTGACAGATCCAGACATGCTCTCCAGAGCTGGATCCAGACTCTGTTATCTTGT 280  
Qy 185 acctgttccatgtacacctgttaacaaagcaacgcgcgtcactcgttggattgggaat 244  
Db 281 ACCTGCTTCATGTCACCTGTAAACCAAGAACCGCTGCTGCTGTGATTTGGAAAT 340  
Qy 245 tcaaacctcttggacatcttgcgcctgagcttggagagcttgaacattacatcatc 304  
Db 341 TCAAACTCTCTGGACATCTTGGCCGAGCTTGGGAAGCTTGAACATTATACATATCTA 400  
Qy 305 gagcttcaaaaaaacaacatccaaagaaactacatcccttgaacttgaagaatcgaagaat 364  
Db 401 GAGCTTCAAAAAACAACATCCAAAGAACTATACCTTCGAACTTGGAAACTGAAGAAAT 460  
Qy 365 ctcatcagcttgcgtctgtacaacaacaactctcaagagatagttcccaactcttggga 424  
Db 461 CTGATCAGCTTGATCTGTACAAACAATCTTACAGGAGATAGTCCACTCTTTGGGA 520  
Qy 425 aaattgaagctctgtgtcttcttcaaggcttaataagacaacggatgagcgggcaatccct 484  
Db 521 AAATTGAAGCTCTGTGCTTTTACGCTTAATGAAACAACGATTCACGCGCAATCCCT 580  
Qy 485 agagcactcaatgcaatcccaagccttaaaagtgttgaatgtctcaagaatgattgtgt 544  
Db 581 AGACACTCAAGCGCAATCCCAAGCCTTAAAGTTGTGAGCTCAGCAATGATTTGTGT 640  
Qy 545 ggaacaatcccaacaagaagaccttctgcacatcccttcaagaacttgaagaacac 604  
Db 641 GGAACAATCCCAACAAGACGACCTTTGCTCACATTCCTTACAGAACTTTGAGAAACAAC 700  
Qy 605 ccgaggttggagggagccggaattactcgttcttgaagctcgaagcaactacgaactga 664  
Db 701 CCGAGATTGGAAGGACCGGAATTACTCGGTCTTGCAGAGCTACGACATACTGCACCTGA 760  
Qy 665 aaaaattggcaaaaacctgaaatgaagaatttgggggtgacacctgttaagaacacttacc 724  
Db 761 AACAACTGGCAAAACCTGAAATGAGAAATGGGGGTGACCTTGTAAAGAACACTTCAC 820  
Qy 725 accttacaatatcacatctactatgttaataagtatataatgtagtccaaaaaataa 784  
Db 821 ACTTATCAAAATATCACATCTACTATATTAATGATATATGATGATGATGATGATGAT 880  
Qy 785 aa 786  
Db 881 AA 882

RESULT 4  
LOCUS A67817 1106 bp DNA PAT 05-MAY-1999  
DEFINITION Sequence 22 from Patent WO9743427.  
ACCESSION A67817  
VERSION A67817.1 GI:4756639  
KEYWORDS  
SOURCE .  
ORGANISM unidentified.  
unidentified.



unclassified:  
1 (bases 1 to 1106)  
De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
AUTHORS  
TITLE  
PRODUCTION OF APOMICRIC SEED  
JOURNAL  
Patent: WO 97/3427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)

BASE COUNT	331 a	258 c	206 g	311 t
ORIGIN				

Query Match	90.5%;	Score 714.2;	DB 5;	Length 1106;
Best Local Similarity	96.2%;	Pred. No. 5.1e-194;		
Matches 753; Conservative	0;	Mismatches 28;	Indels 2;	Gaps 2;

OY	5	ccacgcgctccggagaaacatacggctggggagcccttcgcgagctcgttaacccaaacctta	64
Db	139	CACATGGCGCTCGAACAATCGGTGGGAGCCTTCGAGGCTGGTAAACCTTAACCTTA	198
OY	65	gctttgataccttggttcgaagcaaacctccggaagagatgtcttttcgcgtcttcgcg	124
Db	199	GCTTTGATTAACCTGCTGTCGAGAACCAACTCCGGAAGAGATGCTCTGTAAGCTCTCGCCGG	258
OY	125	agtttaacagatccgagccaatgtctccgaagcctggatccaactcttgttaacccctgt	184
Db	259	AGTTTGACAAATCCAGACCAATGTCCTCCAGAGCTGGGATCCAACTCTGTTAAACCTTG	318
OY	185	acctggttcacatgacacttlaaccaagaacaacgcgcgaactcgtgtgtaatttggggaat	244
Db	319	ACGTGTTCCATGTCACCTGTAACCAAGACAACCGCGATCACGTGCTGATTTGGGAAT	378
OY	245	tcaaaccttccttggaacatctctgcgctggcttgggaagcttggaaactttacatctcta	304
Db	379	TCAAACTCTCTGGACATCTTGCGCTGAGCTTGGAGACTTGAACATTTCACATATCTA	438
OY	305	gagctctacaaaaaacaacacccaaaggaaactataccttcggaacttggaaacttgaagat	364
Db	439	GAGCTCTACAAAACACATCCAAAGAACTATACCTCGCAACTGGAAACTGAANAAT	498
OY	355	ctctcagcttggatctgctgtacaaacaacatcttaaggagatagttcccaactcttcggga	424
Db	489	CTCATCACTTGGATCTGTGTCAAAACAACATCTTACAGGATATGTTCCACTTTCTTGGGA	558
OY	425	aaattgaagctctcgtgctctttttaaagcgttaabgacaacggaattgaaggggccaatccct	484
Db	559	AAATTGAAGTCTCGGTCTTTTAAAGGCTTTAAAGCAACCGATTGACCGGCTCAAT-CT	617
OY	485	agaagacacacttgtaaatcccaagcc-ttaaagtgtgtgactgaagcaatgattgtg	543
Db	618	AGAGCACACACGGCAATCCCAAGCCTTTAAAGTGTGTGACGTCTAAAGCATGATTTGTG	677
OY	544	tggaaacaatcccaacaacaaacggaaacctttgctcaatcctttaaagaaacttgaagaca	603
Db	678	TGGGACATCCCAACAACGAGACCCCTTGTCTCAATTCCTTTACAGAACTTTGAGAACAA	737
OY	604	cccgaggttggaggagaccggaataactcggctcttgtaaactcggacaactaaactgacctg	663
Db	738	CCCGAGATTGGAGGGACCGGAATTACTACGCTCTGTGAACACTACGACACTAATGTAACCTG	797
OY	664	aaaaaaattggcaaaacctgaaaatgaaagaaattggggggttgacctgtlaagaaacttaac	723

[illegible]

	RESULT	5							05-MAY-1998
	LOCUS	A67823						PAT	
	DEFINITION	A67823	Sequence .28 from Patent WO9743427.						
	ACCESSION	A67823							
	VERSION	A67823.1	GI:4756645						
	KEYWORDS	.							
	SOURCE	unidentified.							
	ORGANISM	unclassified							
	REFERENCE	1 (bases 1 to 894)							
	AUTHORS	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.							
	TITLE	PRODUCTION OF APOMICRITIC SEED							
	JOURNAL	Patent: WO 97A3427-A	20-NOV-1997;						
	FEATURES	CTBA GEICY AG (CH)							
		Location/Qualifiers							

BASE COUNT	270 a	163 c	176 g	283
ORIGIN				

Query Match	45.8%;	Score 361;	DB 5;	Length 894;
Best Local Similarity	72.8%;	Pred. No. 7.8e-93;		
Matches 538; Conservative		0; Mismatches 110;	Indels 91;	Gaps 2.

QY 91 ctccgaaggagatbctctcttaacgctcttcgcccgaagttcaacaatccggaccctgctc 150

Db 15 CTCGGAAGGGAGCGCTTCTCAGCGGCTTCCGGAGCTTATCAGATCCAGACAAATGTTGT 74

QY 131 ccgaagcttgagatccaactcttgctatcctctgtatccgctggttcgaatgcaacctgtaacca 210

Db 75 TCAGAGTTGGAGATCCAACTCTTGTTATCTCTGTACTGTGGTCTTGTTATGTGATCATCTTATATCA 134

QY 211 agacacaccgcgtcacctcgttgtagatttggggaattcaaacctctcttgacatctctgcgc 270

Db 135 ACACATCAATCTACTCGTCTGGATTGGGGAATTCAAACTATCTGTGACATCTAGATACC 194

QY 271 tgaagcttgggaagcttgaacatttaagaatat----- 304

Db 195 TGAATCTGGGAAGCTTGAACATTTCAAATATCTGTATGAGATCATCACTCTTTGGCCTT 254

QY 304 ----- 304

Db 255 TGAATATCTGAAAACATTATACATTATCACTCAACACATATTAACATTTTGCTTGGATCAT 314

QY 304 ---agaagcttcaaaaaaacaacatccaagaagaacttaacttcggaacttggaaatctcgaa 360

Db 315 TAGTGAATCTTCAAAAAACAGATTTGAAGAACTATACCTTGTAGAGCTTGGAAATCTGAA 374



QY	361	gaacccatcagcttgatcctcttacaacaaacattcaaggatagctccacattc	420
Db	375	gaGCTTAATCAGTTTGATCTGACAAACAATCTCACCGGGAAATCCACTTCTTT	434
QY	421	gggnaaatgaagctcctgctgctctttacggcttaatgacaacggatgacggccaat	480
Db	435	GGGAAATTAATGATCAGTCTGTTTCTTTCGGCTTAACGAAACCGAATGACCGTCTTAT	494
QY	481	cccttaaggagctacacgacatcccaagccttaagttggttagtgcctcaaggatgatt	540
Db	495	TCTTAGAGACTACAGTTATTCTAAAGCCTTTAAATTTTATATCTCAAGGAATATTT	554
QY	541	gtgtgacaatcccaacaaacgacccctttgctcacatcctcttaacagatcttagaa	600
Db	555	GTGTGACAATTCAGTGAAGGACCTTTGACACATCTCTATGCAAACTTTGAGAA	614
QY	601	caacccaggttggaggagacgggaattctgcgtctgcaagctacagacctaagcac	660
Db	615	CAACCTAGATTGGAAGGACCGAAGCTACTAGCTTTCAGAGCTATGACACCAATTCAC	674
QY	661	ctgnaaaaattggcaaaactgaaatgaagaattgggggtgacacctgtgaagaacct	720
Db	675	TTAAAAAGAGTTGAGAAACCTATAAGAGAA-TGTTAGGTGACCTTGAAAGACTCTG	733
QY	721	caccacttaacaatatc	739
Db	734	TACCAAGTCTTTGTAATC	752
RESULT	6		
LOCUS	A67827	2089 bp	DNA
DEFINITION	Sequence 32 from Patent W09743427.	PAT	05-MAY-1999
ACCESSION	A67827		
VERSION	A67827.1	GI:4756649	
KEYWORDS			
SOURCE			
ORGANISM	thale cress.		
	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
	Arabidopsis		
REFERENCE	1 (bases 1 to 2089)		
AUTHORS	De V.S., Schmidt, R.D., Van, H.G. and Hecht, V.F.		
TITLE	PRODUCTION OF APOMITIC SEED		
JOURNAL	Patent: WO 9743427-A 20-Nov-1997;		
	CIBA GEIGY AG (CH)		
FEATURES	Location/Qualifiers		
source	1..2089		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
	/clone="SEK GENE CDNA"		
	195..2072		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAB42254.1"		
	/db_xref="GI:4756650"		
	/translation="MESSYVVFILLSLILPNHSITLNASNLEGALHTLPTLVDP		
	NILOSDEPLVAPCTWPHVTCNNENSVTRIDDSGLHSLPDELGTLKNOYLETYS		
	NTIGTIPSNLGNLTNLVSDLYNSFSGLTIPSLDGLSLRFLRNNSLTGSLTPMS		
	LNIITLQVLDLSNNRLSGVSDNGSSSLTPTSFANNLDLGPVTSHPCPSPFS		
	PPFIDLPVPSTPSGCTGCAIAGVAAGALPFPAALAFAMWRRSPLDFFEPVSP		
	EEDEPETHGLQKRFSLRELQVADSGFSNKNILRSGFGKYLKGLADGLVAVKLR		
	ERPPGELOFOTREVEIMAVHNRNLRLGCPMPTERLLYVPMANGSVACLERG		
	PSPPLDMPTRIRIALGSAARGSLYLDHCPKIIHRDYKANTILDEEFVAADFGIG		
	AKMDYKDHHTTAVRGTHIAPEVTSKSSSEKIDVGYGIMLELITGGRAVDL		
	RLANDDMLTDMYGLKREKLEMLDPPDLQTYREBELDYIYVALICTGSPME		
	PKSEYVRMLDEGDGLAEKMDKQVEILIREIDLSPPNSDWILDSITLHAEVLG		
BASE COUNT	569 a	419 c	502 g
ORIGIN			599 t

Query Match	26.8%	Score 211.8	DB 5	Length 2089
Best Local Similarity	63.0%	Pred. No. 4.6e-50		
Matches 327	Conservative	0	Mismatches 192	Indels 0
			Gaps	0
Qy	86	gcaactccgaagagatgctctttagctcttcgcgsgagttaacagatccggaccat	145	
Dy	270	GCTAATTTGGAGGAGTGATGCTTGCATCTTTGAGGGTCTACTGATGTCCTCAACAT	329	
Qy	146	gtctccagagcttggaatccactcttctaactcttgtaaccctggttcacatgctac	205	
Dy	330	GTTTCGACAGCTGGGATCCTACGTAGTATCTTGACATGTTCCATGTCACTTGC	389	
Qy	206	aaccaagacaacgcgcctcgtctggtgatttggysgaattcaaacctctctgacatt	265	
Dy	390	AACACGAGACAGTGTCTATAGATGTTATTTGGGAATGACAGATTATCGCCATTTA	449	
Qy	266	ggccctgagcttggaagcttgaacatttaagatcttagagctctcaaacacacac	325	
Dy	450	GTTCCAGAGCTTGCTGTCTCAAGATTTGCAAGTATTTGAGCTTTACATTAACA	509	
Qy	326	caaggaactatccttcggaacttggaaacttgaagaatctcatcagcttgatctg	385	
Dy	510	ACTGCCCCGATTCCTGATATCTTTGGAATCTGACAACTTAGTGAATTTGATTT	569	
Qy	386	aacacacattcaagagtagttcccaactcttttggysaaatgaagctctgycatt	445	
Dy	570	TTAACAGCTCTCCGCTCTCTATTCGGATTCATTTGGGAAAGCTTTAAAGCTGA	629	
Qy	446	ttacgcttaatgaacacggaattgacggggccaaatcccttagagcaactactg	505	
Dy	630	CTCCGGCTTAAACACACAGTCTCACTGGTCAATTCCTATGTCACAGCAATATTA	689	
Qy	506	agacataattgtatgtctcgaagcaatgatttgttgaacaatcccaacaacga	565	
Dy	690	ACCTTCAGATGTATATCTATCAATTAACAGACTCTCTGTTCATGTTCTGCAAT	749	
Qy	566	ctttgtctcaatcccttaccagaacttgaagaacac	604	
Dy	750	TCTCTTCACCTTTCACACCACATGTTTGTCTATTAAC	788	
RESULT	7			
LOCUS	SBU62279	936 bp	mRNA	PLN
DEFINITION	Sorghum bicolor leucine-rich repeat-containing extracellular			
ACCESSION	662279			
VERSION	062279.1	GI:1710123		
KEYWORDS	. sorghum.			
SOURCE	Sorghum bicolor			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.			
REFERENCE	1 (bases 1 to 936)			
AUTHORS	Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.			
TITLE	Isolation of a cDNA encoding a novel leucine-rich repeat motif from Sorghum bicolor inoculated with fungi			
JOURNAL	Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)			
MEDLINE	97124217			
REMARK	Erratum: ([published erratum appears in Mol Plant Microbe Interact 1997 Mar;10(2):302])			
REFERENCE	2 (bases 1 to 936)			
AUTHORS	Hipskind,J.D.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA			
FEATURES	location/Qualifiers			
SOURCE	1..936			
	/organism="Sorghum bicolor"			
	/db_xref="taxon:4558"			
	/clone="SLR"			
	12..755			



[illegible]

REFERENCE	97313247
2 (bases 1 to 1755)	
AUTHORS	Schmidt, E.D.L., Guzzo, F., Toonen, M.A.J. and de Vries, S.C.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-1997) Molecular Biology, Agricultural University of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
FEATURES	<p>source</p> <p>1.. 1755</p> <p>/organism="Daucus carota"</p> <p>/db_xref="taxon:4039"</p> <p>34.. 1665</p> <p>/note="SERK"</p> <p>/codon_start=1</p> <p>/product="somatic embryogenesis receptor-like kinase"</p> <p>/protein_id="AAB61708.1"</p> <p>/db_xref="GI:22249118"</p> <p>/translation="MNRSTNIIILNYMOTFAYILDKYVIMLTLEYSNNISGPIPSDLNITNLSIDLTMNSFSGPIIPDILCKLRLRPLRLNNSSLSGPIPSLNTITLQVLDNNRISGVPDNGSFSLETPISFANNNLILCPVATGRNCSSPPSPSPPIPTSTVOPONPGTGAIGVAAAGALFLFAAPAAAFMMWRARRKREHFEPVPEEPEVHLGQLRSLSELQVADTDEFTIILGRGFGKVKRGLAGSLVAVRLKEEETPGSELQFOTERKEMISAIVRHLRLRGFCMPTERTLIVYPYMANGVASCRLREQPEEPADWETRRKALGSAVRGLIHLHDQDPLIIRQVKANIIILDEEFVAVGDFELALMDYKDHVHTVAGTIGYIAPELISLTKSSEKTDYFGCINILITLTCQSPMERPKMEYVRLEEDSKLLEKDEMLVDLDENNNTIDTVEKQLQVALLCTQGSPEKPRPKMEYVRLEEDIAEKDEWQKVEVHODELAPHTPTSSWILDSTDNCLAFELSGPR"</p>
BASE COUNT	506 a 347 c 407 g 495 t
ORIGIN	
Query Match	13.1%; Score 103.4; DB: 8; Length 1755;
Best Local Similarity	55.4%; Pred. No. 5.2e-19;
Matches 200; Conservative	0; Mismatches 161; Indels 0; Gaps 0
Qy	265 tgcgcctgagctggggaagcttgaaacattacagatctagaagctctacaacaaacacat 324
Db	78 TGAGCTTACCTTGACAAATATGGGGTCTCTTATGACATTGGAGCTTACACAAATACAT 137
Qy	325 ccaaggaactataccttcgcgaacttggaaattgagaatctcaatcagcttgatctgta 384
Db	138 AAGTGGACCAATTCCTAGTGAATCTTGCGAATCTGCAAAATTGGGAGCTTGAGCTATA 197
Qy	385 caaacaacatcttacagggaagtagttcccaactcttggggaataatggaagtcctgctt 444
Db	198 CATGAATAGCTTCTCTGGACCTTACCGGACACATTAGGAAGCTTACAAGCTTAAGATT 257
Qy	445 ttacagcttaatgacaacagcattgacggggacatccctagagcactacatgcaatcc 504
Db	258 CTTGCGTCTCAACAACAGCCTCTCTGTCCTCAATTCGCAATGCTCAAGCTCAATATATAC 317
Qy	505 aagcctaaagttgttgatgctcacaagcaatgatttggttggaacaatcccaacaaagg 564
Db	318 AACTCTTCAAGTCTCTGGATTATCAACAATGGCGCTATAGAGCCAGTACCGGATATGG 377
Qy	565 acccttgcacatctcctttacagaacttgaagacaacccgaaggttgagggagccga 624
Db	378 CTCATTTTCTTTGTTATACACCTATACGTTTTCGCAATATATTGAATTTATGAGCCGT 437
Qy	625 a 625
Db	438 A 438
RESULT	9
LOCUS	A67797 1815 bp DNA PAT 05-MAY-1999
DEFINITION	Sequence 2 from Patent WO9743427.
ACCESSION	A67797
VERSION	A67797.1 GI:4756623
KEYWORDS	carrot.
SOURCE	Daucus carota
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.  
REFERENCE  
1 (bases 1 to 1815)  
AUTHORS  
De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE  
PRODUCTION OF APOMICRITIC SEED  
JOURNAL  
Patent: WO 974327-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
Location/Qualifiers  
1..1815  
/organism="Daucus carota"  
/db\_xref="taxon:4039"  
94..1755  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42241.1"  
/db\_xref="GI:4756624"  
/translation="MNRSINILNTMQFTDAYLDKRYGLTMELEYSNISGPIPSDGL  
NLNIVSLDLYNNSFSGPIPDYLGRLTLRLRLNNLSGPIPSLITTLQVLDL  
SNRLSGPDPNSFSLFPIPSFANNLNCGVYGRPCGSPSPSPPEFIPSTVOP  
PGONGPTGAAGGVAAGALFLPADAMAFWMRRKPREHFDVPAEDRPHVHGLK  
RSLRELOYATDPTSTIILRGSGKGVYGRGLADGSLVYKLEKRRPGSLQOTVY  
EMISNAYHRNLRLKRFCTPTPERLIVPYMANGSVASCLREOPSEPPDPTKRL  
ALGSARGLSYLDHDCPKIHRDVKANLILDEFEAVVGDGLARLDYKDTHTTA  
VRGTLGIAPERYLSTGSESEKTDVFGIMLDELITGQAPDLARLANDVDMLDLY  
KSLKEKKLEMLVDPDENNYIDTEVEQLIYALLCTQSGSPMERKMESEVYRMLECDG  
LAEKMEMQKVEYIHODVELAPRHSSEMLDSTDNLHAFELSGPR"  
BASE COUNT 531 a 354 c 415 g 515 t  
ORIGIN  
Query Match 13.1%; Score 103.4; DB 5; Length 1815;  
Best Local Similarity 55.4%; Pred. No. 5.2e-19;  
Matches 200; Conservative 0; Mismatches 161; Indels 0; Gaps 0;  
QY 265 tgcgcctgagcttggagagcttgaacattacagttatctagagcttacaacaacacat 324  
Db 138 TGAGCTTACCTTGACAAATATGGGCTTATGACATTGACCTTACACCAATTAACAT 197  
QY 325 ccaaggaactatactcgaacttggaaatcgaagaatcctcatcagcttgcattcga 384  
Db 198 AAGTGACCAATTCCTAGTATCTTGGAATCTGACAAATTTGGTAGCTTGACCTANA 257  
QY 385 caacaacaacttacaagagtagttccacttcttgggaaattgaattgcctcctgctt 444  
Db 258 CATGATATGCTTCTCTGACCTATACCGACAAATTAGGAAGCTTACAGGCTAAGATT 317  
QY 445 ttacggcttaaatgacacacgattgaacgggccaatccctagaacacatcgaatccc 504  
Db 318 CTTCGCTCTGACAAACAAGCCTCTCTGCTCAATTCGAATCTCACTGACTAATATTAC 377  
QY 505 aagccttaagcttgcattgtctcaagaatgattgtgtggaacaatcccaacaacgg 564  
Db 378 AACCTTCAAGTCTGATATATCAAAACATCGGCTATCGGACGATACCGGATATG 437  
QY 565 acccttgcctacattcctttagaacttggaaacacccgaggttggagagacgga 624  
Db 438 CTCAATTTCTTGTATTACACTATCACTTTGCCAATAATTGAATTATATGACCCGT 497  
QY 625 a 625  
Db 498 A 498  
RESULT 10  
LEIRGENE  
LOCUS 4604 bp DNA PLN 10-OCT-1996  
DEFINITION  
L.esculentum LRP gene.  
ACCESSION  
X95269.1 GI:1619299  
VERSION  
X95269.1 GI:1619299  
KEYWORDS  
LRP gene; LRP protein.  
SOURCE  
tomato.  
ORGANISM  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
Asteridae; Solanaceae; Solanales; Solanales; Solanaceae; Solanum; Potatloe;  
Lycopersicon.  
REFERENCE  
1 (bases 1 to 4604)  
AUTHORS  
Vera,P.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (18-JUN-1996) P. Vera, Universidad Politecnica de  
Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino  
de Vera 14, E-46022 Valencia, SPAIN  
2 (bases 1 to 4604)  
AUTHORS  
Torneiro,P., Mayda,E., Gomez,M.D., Canas,L., Conejero,V. and Vera,P.  
TITLE  
Characterization of LRP, a leucine-rich repeat (LRP) protein from  
tomato plants that is processed during pathogenesis  
JOURNAL  
Plant J 10 (2), 315-330 (1996)  
MEDLINE  
9636767  
FEATURES  
Location/Qualifiers  
1..4604  
/organism="Lycopersicon esculentum"  
/cultivar="VF8"  
/db\_xref="taxon:4081"  
/tissue\_type="leaf"  
1..345  
/number=1  
113..4355  
/gene="LRP"  
join(113..345,2381..2452,2565..2708,2928..2999,3081..3152,  
4283..4355)  
/gene="LRP"  
/codon\_start=1  
/product="LRP protein"  
/protein\_id="CAA64565.1"  
/db\_xref="GI:1619300"  
/db\_xref="SPTREMBL:O96477"  
/translation="MEAVEKQSLFLKIMGLAVYLAVAVVAGNSRQDALVALRSL  
SPRGVLOSMDPNIYNPCTWTHVYCNQNDQYTRDLNLSKLSHLYPELKLHLYDL  
EIKNNIGTILPEKLGKSLISLDLNNISGIIPTSLGKLVFRLUNDKLTGP  
IRRELTSISLKVVDVSNNDLCGIIPTSGPEHILNNFHNPRLEGPILLGATYDT  
NCS"  
346..2380  
/gene="LRP"  
/number=1  
2381..2452  
/gene="LRP"  
/number=2  
2453..2564  
/gene="LRP"  
/number=2  
2565..2708  
/gene="LRP"  
/number=3  
2709..2927  
/gene="LRP"  
/number=3  
2928..2999  
/gene="LRP"  
/number=4  
3000..3080  
/gene="LRP"  
/number=4  
3081..3152  
/gene="LRP"  
/number=5  
3153..4282  
/gene="LRP"  
/number=5  
4283..4604  
/number=6  
BASE COUNT 1311 a 758 c 888 g 1647 t  
ORIGIN  
Query Match 12.5%; Score 99; DB 7; Length 4604;  
Best Local Similarity 77.4%; Pred. No. 1e-17;



Matches	120;	Conservative	0;	Mismatches	35;	Indels	0;	Gaps	0;
QY	80	gtcgaagcaactcgcgaagagatgctcttaacgctcttcgcggaggttaacagatccg	139						
Db	194	GTAAAGGGGAAATTCACAAAGGGGATGTTGTTATAGCCCTTCGCGAGCTTATGTGACCCG	233						
QY	140	gaccatcttcctcagagctcgagatcccaactctgttaactcttacctggtttccatgctc	199						
Db	254	GGTACTGCTGTAAAGAGCTGGGATCCAAATCTTGTTAACTTGTACTCTGGTTTCATGTC	313						
QY	200	acctgttaaccaagacacgcgcgtactcgtgtgga	234						
Db	314	ACTTGCAACGAGATATCATCAAGTTACTCGTGCTA	348						
RESULT	11								
ATAC009991/c									
LOCUS	ATAC009991	101284 bp	DNA	PLN	08-OCT-1999				
DEFINITION	Arabidopsis thaliana chromosome III BAC F9P8								
ACCESSION	AC009991								
VERSION	AC009991.3								
KEYWORDS	HTG.								
SOURCE	Arabidopsis thaliana								
ORGANISM	Arabidopsis thaliana								
REFERENCE	1 (bases 1 to 101284)								
AUTHORS	Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B., Ronning,C.M., Koo,H., Fujii,C.Y., Utecherack,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.								
TITLE	Arabidopsis thaliana chromosome III BAC F9P8 genomic sequence								
JOURNAL	Unpublished								
REFERENCE	2 (bases 1 to 101284)								
AUTHORS	Lin,X. and Kaul,S.								
TITLE	Direct Submission								
JOURNAL	Submitted (09-SEP-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org								
REFERENCE	3 (bases 1 to 101284)								
AUTHORS	Lin,X.								
TITLE	Direct Submission								
JOURNAL	Submitted (08-OCT-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA								
COMMENT	On Oct 8, 1999 this sequence version replaced g1:5902413. Address all correspondence to: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA e-mail: xlin@tigr.org BAC clone F9P8 is from Arabidopsis chromosome III and is near the molecular marker g4547. The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of three methods: Gene Prediction programs including GRATU (available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), GenScan (Chris Burge, http://genomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene (http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (http://www.tigr.org/cdb/at/est.html). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins Genes without protein or EST similarity, that are predicted by morr than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are								

identified by RepeatMasker (url: <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAT are annotated as misc features.

Location/Qualifiers

```
1. .101284
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="11"
/map="q4547"
/clone="F9F8"
77. .108
/misc_feature
/note="exon predicted by xgrall, quality excellent"
124. .259
/misc_feature
/note="exon predicted by xgrall, quality excellent"
complement(347..408)
/misc_feature
/note="exon predicted by xgrall, quality
excellent_shadowexon"
355. .594
/misc_feature
/note="exon predicted by xgrall, quality excellent"
672. .752
/misc_feature
/note="exon predicted by xgrall, quality excellent"
854. .973
/misc_feature
/note="exon predicted by xgrall, quality excellent"
complement(1182..1357)
/misc_feature
/note="exon predicted by xgrall, quality
marginal_shadowexon"
complement(1516..1694)
/misc_feature
/note="exon predicted by xgrall, quality
excellent_shadowexon"
1546. .1666
/misc_feature
/note="exon predicted by xgrall, quality good"
1747. .1838
/misc_feature
/note="exon predicted by xgrall, quality excellent"
complement(2090..2250)
/misc_feature
/note="exon predicted by xgrall, quality good_shadowexon"
2114. .2246
/misc_feature
/note="exon predicted by xgrall, quality excellent"
complement(2975..3024)
/rpt_family="CNA)n"
/rna
join(3070..3325,3565..3642,4094..4234,4338..4414,
4506..4655,4800..5013)
/gene="P9F8.1"
3070. .5013
/gene="P9F8.1"
/note="similar to unknown protein GB:AAC37167 [Arabidopsis
thaliana]"
join(3215..3325,3565..3642,4094..4234,4338..4414,
4506..4655,4800..5013)
/gene="P9F8.1"
/note="unknown protein"
/codon_start=1
/protein_id="AAFO1505.1"
/db_xref="gi:6016678"
/translacion="MWGAPQIVFGSSIVOMRSGHGMGAILSEVYARRAKDITLR
KYNAGSRALVEVDGPFKDAVOPSLIVYEGNDNAPSSGLGPPVPLEYVDNN
KRIAHHSISDFRIRIFLSRPNDEAVRONQSYLSEVIRTDNICKYDACEVLE
QELGLEAVDLESTFOKADMDKIVCTDGIHLSAOSKTYVAGEILRYKAEALHPSLH
KSMPEFADSDSPYDILVSDGQVNSSEMTYFWEKQW"
complement(5513..5597)
/rpt_family="GA)n"
/complement(5527..5597)
/note="exon predicted by xgrall, quality
excellent_shadowexon"
join(5589..5795,5897..5929,6013..6235,7264..7387,
7467..7932)
/gene="P9F8.2"
5589. .7932
/gene="P9F8.2"
/note="similar to nucleic acid binding protein Alfin-1
GB:AAA20093 [Medicago sativa]"
join(5696..5795,5897..5929,6013..6235,7264..7387,
```



Query Match	10.7%	Score 84.8	DB 8	Length 101284
repeat_region	1467. .7727			
repeat_region	/gene="p9F8.2"			
repeat_region	/codon_start=-1			
repeat_region	/product="putative nucleic acid binding protein"			
repeat_region	/protein_id="AAF01506.1"			
repeat_region	/db_xref="GI:6016679"			
repeat_region	/translation="MAAAVSSNPRTVEEIEFKDYSARRAALLRLATKTDVDYFSCDPP			
repeat_region	KEENCLCYHPHSEWEVNLPAEEYPELPEPPLAGIFADGKQMDLTLVAVHSDD			
repeat_region	LEVSIFYEARALNRERKRLFLSLINDPLFTDVTGGRKMKQNRSSDSGSKSNQK			
repeat_region	RSIDQGTSSPMTKMESEYEEDEEDHGDTLGGSGGHTTHEFWICDVCENYHSG			
repeat_region	KCYKITPAKASEIKQYKCPCCAKKQ"			
repeat_region	complement(771. .7760)			
repeat_region	/rpt_family="(CAT)n"			
repeat_region	complement(7810. .7847)			
repeat_region	/rpt_family="(GAAA)n"			
repeat_region	8515. .8572			
repeat_region	/rpt_family="(TAA)n"			
repeat_region	complement(join(<9426. .9674,10198. .10528,11098. .11345,			
repeat_region	11491. >11865))			
repeat_region	/gene="p9F8.3"			
repeat_region	complement(<9426. >11865)			
repeat_region	/gene="p9F8.3"			
repeat_region	/note="similar to leucoanthocyanidin dioxygenase			
repeat_region	GB:BA420143 [Perilla frutescens]"			
repeat_region	complement(join(9426. .9674,10198. .10528,11098. .11345,			
repeat_region	11491. .11865))			
repeat_region	/gene="p9F8.3"			
repeat_region	/codon_start=-1			
repeat_region	/product="putative leucoanthocyanidin dioxygenase"			
repeat_region	/protein_id="AAF01507.1"			
repeat_region	/db_xref="GI:6016680"			
repeat_region	/translation="ANNLDIEIKSTCTCLNDQOEYKINDMHMSDQDKNKIEIKNSG			
repeat_region	LGKMPEDIVRVQSLAESNLSPDRYIKPPSORQTTIIDQPVADINIDPILDLS			
repeat_region	LEFGNDQDKRKISEACREMGFPYQYINIGVKEPMADARFMSFNLVEAKEYSSKL			
repeat_region	PRYEEYSGRSLVEKGATLIDNDYDYLAFPLAKQNPWPGLSIEENMDYDEKRL			
repeat_region	YKIGGLMTILSSNLGLARQLOAFGEDEVGACLNRYPPCPPELALGLSPHSPP			
repeat_region	GGATILPDQYVGLQVRKGPWLTVPNPLKATVINGDQIQLNSKSKVERIVAT			
repeat_region	NSKEKESVLAFFINPKRSIDPIPIQMOQLVSTMPPLVPMTFDQYRLFIRTGQPGKSH			
repeat_region	VESHKIPR"			
repeat_region	11962. .11996			
repeat_region	/rpt_family="(GA)n"			
repeat_region	12824. .12836			
repeat_region	/note="exon predicted by xgrall, quality			
repeat_region	marginal.shadowexon"			
repeat_region	complement(13367. .13468)			
repeat_region	/note="exon predicted by xgrall, quality marginal"			
repeat_region	complement(13742. .13788)			
repeat_region	/note="exon predicted by xgrall, quality marginal"			
repeat_region	complement(join(13822. .14179,14261. .14398,14502. .14582,			
repeat_region	14676. .14861,14941. .15033,15124. .15190,15272. .15361,			
repeat_region	15662. .16380))			
repeat_region	/gene="p9F8.4"			
repeat_region	complement(13822. .16380)			
repeat_region	/gene="p9F8.4"			
repeat_region	/note="identical to omega-3 fatty acid desaturase,			
repeat_region	chloroplast precursor GB:P46310 [Chloroplast Arabidopsis			
repeat_region	thaliana]"			
repeat_region	complement(join(13988. .14179,14261. .14398,14502. .14582,			
repeat_region	14676. .14861,14941. .15033,15124. .15190,15272. .15361,			
repeat_region	15662. .16155))			
repeat_region	/gene="p9F8.4"			
repeat_region	/codon_start=-1			
repeat_region	/product="omega-3 fatty acid desaturase, chloroplast			
repeat_region	precursor			
repeat_region	/protein_id="AAF01508.1"			
repeat_region	/db_xref="GI:6016681"			
repeat_region	/translation="MANVILSECGIRPRLRYTPPNSNFKLNKKRPSLSSSYKTS			
repeat_region	SSPISFLNRDGFETNNALVNSPFLTPPISESPLEEDKQDPGAPPPNADIRIS			
repeat_region	AAIDKHQGNPKPKSLSYVRDVAIVAFVLAAGAVLNNTWVPIVTLAOGTAPATVAT			
repeat_region	LGDCGCGSSNDPRKINSYGVILHSSITLYDHGRKRSSETHQNGHYENDESHRP			
repeat_region	SEKTIINDLDFEERFFQLPLVMTALPIPTLWANSFGKSGSHIPDSDLFLPERADVA			

Matches	167	Conservative	0	Mismatches	137	Indels	0	Gaps	0
Best Local Similarity	54.9%	Pred. No. 1.6e-13							
QY	252	tctcttgacatcttcgcgcctgagcttggaagcttgaacattacagtcatagcctc	311						
Db	44936	TCATTGGTACAGATTCCATCTTCGATTGGGAAACCTTGCTCTCCACCTATCTTACCTTT	44877						
QY	312	acaaagacacatcccaaggaactacatcttcgcgaacttcgaatctcgaagaatcctaca	371						
Db	44876	CTTACACAACTTTCGTGGTGAATCCCATCTCTCTTTGGCATCTTAAACACACTGATG	44817						
QY	372	gcttggaatctcgtacaaacaaactcttcaggagatagttcccaactcttgggaatctga	431						
Db	44816	TCCTTACAGGTGATTCCTCAATAGCTACATGGAAATGTTCCCATCTCACTCAATTTTA	44757						
QY	432	agctctggctcttttaccagcttaatgacacagcatggaacgggccaactccctaggac	491						
Db	44756	CGAGGTTCCTGCTTATTACTTTCCTCCACAACTCAGTTCACAGGACGATTCCTCAATTA	44697						
QY	492	tcacatcgaatcccaagccttaagtgttgttgctcgaagcaatgatgttgtagaaca	551						
Db	44696	TCAGTTTACTATCAAACTTGATGAGCATTTTGAGCAAGTAACAACTTCCTCAGTAATC	44637						
QY	552	tccc	555						
Db	44636	TCCC	44633						
RESULT	12								
AC008017	116944	bp	DNA	PLN	18-SEP-1999				
LOCUS	Arabidopsis thaliana	chromosome I	BAC F3AN23	genomic sequence,					
ACCESSION	AC008017			complete sequence.					
VERSION	AC008017.2			GI:5487890					
KEYWORDS	HTG.								
SOURCE	thale cress.								
ORGANISM	Arabidopsis thaliana								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core								
	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;								
	Arabidopsis.								
REFERENCE	1 (bases 1 to 116944)								
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,								
	Altai,H., Arujo,R., Hulzar,L., Rowley,D., Chen,S., Herman,P.,								
	Hicks,R., Huerta,M., Mason,S., Stepel,J., Zimmerman,M., Buehler,E.								
	Dunn,P., Gonzalez,A., Kremenskaia,I., Kim,C., Lenz,C., Li,J.,								
	Iiu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaja,V.,								
	Vyotskaja,V.S., Walker,M., Yu,G., Ecker,J., Theologis,A. and								
	Davis,R.W.								
JOURNAL	Unpublished								
REFERENCE	2 (bases 1 to 116944)								
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,								
	Altai,H., Arujo,R., Hulzar,L., Rowley,D., Chen,S., Herman,P.,								
	Hicks,R., Huerta,M., Mason,S., Stepel,J., Zimmerman,M., Buehler,E.								
	Dunn,P., Gonzalez,A., Kremenskaia,I., Kim,C., Lenz,C., Li,J.,								
	Iiu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaja,V.,								
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.								
TITLE	Direct Submission								
JOURNAL	Submitted (10-Jul-1999) DNA Sequencing and Technology Center,								
	Stanford University, 855 California Avenue, Palo Alto, CA 94304,								
REFERENCE	3 (bases 1 to 116944)								
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,								
	Altai,H., Arujo,R., Hulzar,L., Rowley,D., Chen,S., Herman,P.,								
	Hicks,R., Huerta,M., Mason,S., Stepel,J., Zimmerman,M., Buehler,E.								
	Dunn,P., Gonzalez,A., Kremenskaia,I., Kim,C., Lenz,C., Li,J.,								
	Iiu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaja,V.,								



REFERENCE 4 (bases 1 to 116944)  
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altati,H., Araujo,R., Huizar,L., Rowley,D., Brooks,S., Bieher,E.,  
Chao,Q., Dunn,P., Gonzalez,A., Khan,S., Kremenetska,I., Kim,C.,  
Ienz,C., Li,J., Liu,S., Lueros,S., Schwartz,J., Shin,P.,  
Toriumi,M., Vytotskaia,V., Yu,G., Ecker,J., Theologis,A. and  
Davis,R.M.  
TITLE Direct Submission  
JOURNAL Submitted (18-SEP-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
COMMENT On Jul 15, 1999 this sequence version replaced g1:5441915.  
FEATURES  
source location/Qualifiers  
1..116944  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F3N23"  
/complement(133..2966)  
/gene="F3N23.1"  
/complement(join(1333..2271,2355..2966))  
/note="81% identical to threonine synthase [Arabidopsis  
thaliana] (g1:4914408). Location of ests GBGF589 3'  
(gb|F15267) and GBG6175 (gb|226034) "  
/codon\_start=1  
/product="Putative threonine synthase"  
/protein\_id="AAd55628.1"  
/db\_xref="GI:5903070"  
/translation="MASFLPSATYFPSSHSELSLPHSAASFVTCSTASPAVPPQT  
POKPRSPINDEARRRPHOLNLSARVPNAPPSSTESYSLDEIVRSQALL  
DNOHDFALKRYDEFFMRNLFDSRNPAPPSGVSMSKRWLPELIDDDIYRASFEG  
NSNLFPAERGRKOYLOMNDLWYKHCISHRGSKRDLGNSVLRKLRKRPYIG  
GCGSTDTSAALSHYCSAGISTVFLPRADIKSMALVPIANGAFVLSITDIPGCM  
HLIREVTELPITLANSLSRLRGQTALETLLQGFNMQVPIVPGCNGNTAF  
YKGFHMKELGLVDRIPLVCAQANANPLYLKSGFKEDNPILKANTFSAIQG  
DPSIDRAVYALKRNSIVGEATEEELMDATLADSTGMEICPTGTGATLALMKLKS  
GVGANDRTVYVSAHGLKFTQSKIDYHSKNIKEMACRLANPVPVKAKFGSMDVLK  
EYLKSNDR"  
gene 6030..7079  
CDS /gene="F3N23.2"  
6030..7079  
/note="F3N23.2"  
/note="Unknown protein; location of ests 205D12T7  
(gb|H77203) and 205D12XP 3' (gb|AA605559) "  
/codon\_start=1  
/protein\_id="AAd55629.1"  
/db\_xref="GI:5903071"  
/translation="MNLGAEEESAQELHLPADINWMDLRSKFVYLGALFSGVSGA  
LYPAVLNKTROVCHSQSGCIKTAFTLVREHGLRGYRGFTSLMGITIPARALYMAL  
EYKSNVNGSAVSLGLEAKAAVANAAGLSAAMAQLWTVIPDVVSRIMVQSGAG  
LVNARCNVYNGFAPFRKIYRADPGKGLYAGFTGISTLYPASNAMVMSVAVRW  
GGICGYCKRDEESGNTMTPDKSKIMAVQVSALIGSVSALITMDPTDKTFLD  
VLDEDSNNKRSRPSIGQTVRNLYREGGWTACTRGIGPCASMSASATMTTITTEFL  
KRLSAKHHDGYSKS"  
complement(8272..9661)  
/gene="F3N23.3"  
/complement(join(8272..8658,8746..8916,9005..9079,  
9161..9252,9364..9661))  
/gene="F3N23.3"  
/note="98% identical to Transcription factor [Arabidopsis  
thaliana] (g1|2398525) "  
/codon\_start=1  
/product="Transcription Factor"  
/protein\_id="AAd55630.1"  
/db\_xref="GI:5903072"  
/translation="MMHOMLKKKDSATHSPLPNTISMGVYPTDSVANRBSASLS  
SLKVDSPRGHITQTKQISFDQDSSSTQSTGQSTEVASSGDNDPSKISFSKSSSE  
ITQKRGFASNPKQSGMTGFPIHHPAPQANFSFIADPHYGGLAATYVLAQPTCPQO  
MYSIPIGVPLPAELTETDPVYVNAKOYHAIMRRQGRALKEAQNKILIRAKKYLIES  
RHVHALKRRSGGRFLNTKTLQESQAAAREODKLGCOVNRKTKNSRFPAHMQ  
NNKRSSTSSSDITSVDGADIDIGHTEFQPSGPTPIINNALVHQGSDMHGCGDMH  
HPVHTI"

gene complement(12076..16057)  
CDS /gene="F3N23.4"  
complement(join(12076..12519,12613..13851,13955..14140,  
14368..15451,15672..16057))  
/gene="F3N23.4"  
/note="Similar to disease resistance proteins RPP1-WSB  
(g1|3860165), RPP1-WSG (g1|3860167) and RPP1-WSA  
(g1|3860163) [Arabidopsis thaliana]. "  
/codon\_start=1  
/product="Similar to disease resistance proteins"  
/protein\_id="AAd55631.1"  
/db\_xref="GI:5903073"  
/translation="MASSSSSATRLRHVDVFLSPRGVDTROTIVSHLYVALRNGVYL  
TRDPDKLEIGDTIADGIYKATQTSWFAVYVILSENAVSTWGLEELRIMQHSQEI  
KVLPIFGVAPSVYRQEGSFATAPQSYDEADMAVAVGISTRSQPRACFTLENVG  
MEAHMMKMTLLNIGCEDEHYHMLGIMGKGIQSTAKLITRFSKQPRACFTLENVG  
KGYDKHLQKELLSHLYDEYDEVMENMGSEIKRRLHQVQFVLDVNDVEQLHG  
LADPSWFGGSRITITTRDKGLNCSGVNITVEYKCDKDALQEFKRLAFGRPPS  
DGEFOLFIASRLAHGLPSALVFASHLSAIVADMEDELLALEFPQKNQELIIRA  
SYDGLDQYDKTVFLHVAQCPNNGHLITRPAFKNCARLNHAKCLVIVISDGCISM  
HILLVQGRIVYQESDMSRSKRFMDPTEPIHYVDSNTHLGAVNSQLISDDYVL  
SRNKLTHMDATPILTPILFRPHTIIEISLRKSLNSMDGTCLPILRIIDVGRS  
NLRLPELSTAVALLEULLESCTSLVQIPESIRLYRLRLNMYCDGLGVLIVDNLQ  
EASLSRWGLKRIILNPHSGATLSLTDALIQKIPKLSGSGTEGDLNLSFSSVQTA  
HOSVTHLNGSFGFKSLDKRPSRLDPVNECLSFADPCLTELKTLNLTIEDIPE  
DICOJOLLETLIDGNDYVILPTSMQOLAMKLSLNSCRRLKALPOLSOVERLYSG  
CYRGSIMGLTIGAGRYNLDPQCYKSGISGLITISVSKAPGRNELLESLNKS  
LVLSSELSEFTKLTITLDSLSLFRPIPTISIRLSMRTILYNLCKITSLIDPESL  
KYLVAHGESLEHVNFSNHSFNHLDPSHCISLSECLSDVRFPMNFSQDEAPFLVC  
ITKYSITSNMRTSMREPMRILPRIKAPKLVGFVQIMVACEPFLHQPASYN  
WDQGSRLYILINKPILYQSEKMEENNNRPMYKMHVLVIOIPTGITSAEIDVOVES  
HVOYEEPEDEILICGVHGVFLK"  
16543..17907  
/gene="F3N23.5"  
join(16543..16961,17058..17907)  
/gene="F3N23.5"  
/note="Similar to N-terminal half of downy mildew  
resistance protein RPP5 [Arabidopsis thaliana]  
(g1|2109275) "  
/codon\_start=1  
/product="Similar to part of downy mildew resistance  
protein RPP5"  
/protein\_id="AAd55632.1"  
/db\_xref="GI:5903074"  
/translation="MSSSKFEVFLSECSDEPKTYVSLDRLMEQRDITTNKDDSF  
LAESKLAIVVYSESYPIVYLCLNQLKINYSSEGLSTLPFYFQVDPYVNRQGY  
LAPROELREGYDPDKIOEMRVSLTKTLNIPALDSRYMSMEADMTILANEISIR  
KPLAKGDLVGNDRQKQITXKLLDFAAEVYVILGIGWGGICGTTLARAYEETIS  
NFKVHVEDKAEKICHDRODLKLLKRGTKGIDGIDIKIKSTFGHKGILYIDCYD  
NIOKLEIVLAHWFIPGSRYITVQDRNLIVSGVHALEVOSLRTDEALQFSHSA  
FDQOHPPTSESLSLRAVHISGFLPLTLKILIGSSLKGDKEBERKEKLOLDEGOKAI  
MEITSKRTYRAGKEEDKEKITSITLSDS"  
complement(19050..23350)  
/gene="F3N23.6"  
complement(join(19050..19667,19764..21344,21686..22799,  
22887..23350))  
/gene="F3N23.6"  
/note="Similar to entire sequence of downy mildew  
resistance protein RPP5 [Arabidopsis thaliana]  
(g1|2109275) "  
/codon\_start=1  
/product="Similar to downy mildew resistance protein RPP5"  
/protein\_id="AAd55633.1"  
/db\_xref="GI:5903075"  
/translation="MASSSSSPINRWYVFLSFRGEDTRKNIVSLHKLQVLDKGVVTFK  
DKKLEIGDSISEIRAIQNSTALVILSENVASSMGCDELRLMDVMDHLKRIKIV  
PIFYQVDSHYRHQTSRTEBDKQDSSMPKVTWRALQVIAISGKDECEDEAS  
MTEIVKDISKLLIMQVDPDSIVGNAMHERRISPLMSDSENVPMIGKMGIG  
KTTIAKLEDFESQFPAPRCLENVASITRRGVSSIAEFLSTTLGSLKRNKGSV  
KLGPETIARCEGRKRVFVVDNDVDMQMAFQDESSWFQPSRIITTRDKGLNMY  
GVRTYEVKCDNDALALQLEFQALFKALPSELLEYKLSIASMLAQGLVAAEAGL  
FFRRMTSLKEMDNLACRITFAPDSVEIKITISGVEETDKVFLHACLPNGEPLR  
BATITLDGTVAGGLGLIAEKLEIITASGYIKMNLVDOGRARALVNGESQRRRG







RESULT	14
AC021198	LOCUS
AC021198	81513 bp DNA
AC021198	PLN 09-FEB-2000
AC021198	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete sequence.
AC021198	AC021198
AC021198.2	GI:6957696
HTG.	
thale cress.	
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eumphylophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 81513)
	Lin,S.X., Sakano,H., Yu,G., Lee,J.M., Lenz,C., Pham,P., Tortumli,M., Chin,C., Chlou,J., Choi,E., Chung,M., Gonzalez,A., Howng,B., Liu,A., Vaysberg,M., Altafi,H., Brooks,S., Buehler,E., Chao,O., Conn,L., Conway,A.B., Hansen,N.F., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.
TITLE	Arabidopsis thaliana chromosome 1 BAC F14D7 sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 81513)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE	3 (bases 1 to 81513)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	On Feb 11, 2000 this sequence version replaced gi:6693723. The sequence is of BAC F14D7 from Arabidopsis thaliana chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 6954 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide small overlaps (200 bp) between overlapping submitted clones. The 5' end of this sequence overlaps by 200 bp to the 3' end of the sequence of the clone F1504.
FEATURES	Location/Qualifiers
SOURCE	1. 81513
	/organism="Arabidopsis thaliana"
	/cultivar="Columbia"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="F14D7"
BASE COUNT	27779 a 13814 c 13868 g 26052 t
ORIGIN	
Query Match	10.1%; Score 79.8; DB 50; Length 81513;
Best Local Similarity	53.7%; Pzed. No. 4,1e-12;
Matches 165; Conservative	0; Mismatches 142; Indels 0; Gaps 0
249 accctctggagcactctggcgccctgagctggggaagctgaatttaagatcatagacg 308	
Db 2466 ATCTACTGGGGGCGATTCGCCCGAAGCTAGGCAACTAGAACATCCATGATCATTTAGGT 2555	
QY 309 tctacaaaacaaacatccaaaggaactataacctccgaacttggaaattggaatctgaatctca 368	
Db 2526 TGAGTAACAACAACCTTACCGGTTCAATTCCTTCCTTAGGAATCTCAAGAAATTGA 2585	
QY 369 tcagcttgatctctgacacacaaactcttaacaggaatagttcccaactctcttgggaaat 428	
Db 2586 CTATCTTACTTATCTTAGAGAACTATTTAATCTGGTGTCAATTCGCCCGGAAGCTAGCAACA 2645	

Oy	429	tgaagctcctgctcttttaacggcttaataaacacgcattgacggcgccatcccttagag	488
Dd	2646	TGGATTCGATTGATTTACAGTGTGAATAACAACAACACTTACC GGTCGAAATCCCTTCTT	2705
Oy	489	cactactcgtcaatcccagaacctaaagtctgtgatgtctcaagaatgatttcttgaa	548
Dd	2706	CCTTCGCAAAATCAGAACCTTGACGTATCTTTATCTTACCTCGAATATTATTAACCTGTG	2765
Oy	549	caatccc	555
Dd	2766	TCATTC	2772
RESULT	15		
LOCUS	ATAC011765		
DEFINITION	Arabidopsis thaliana chromosome I BAC FIM20 genomic sequence,	PLN	08-DEC-1999
ACCESSION	AC011765		
VERSION	AC011765.4		
KEYWORDS	HTG.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core		
	eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
REFERENCE	Arabidopsis.		
AUTHORS	1 (bases 1 to 134402)		
	Lin,X., Kaul,S., Town,C.D., Beitel,M., Creasy,T.H., Haas,B.,		
	Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,		
	Bowman,C.L., White,O., Niernm,W.C. and Fraser,C.M.		
JOURNAL	Arabidopsis thaliana chromosome I BAC FIM20 genomic sequence		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 134402)		
	Lin,X. and Kaul,S.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (14-OCT-1999) The Institute for Genomic Research, 9712		
AUTHORS	Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org		
	3 (bases 1 to 134402)		
JOURNAL	Lip,X.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (08-DEC-1999) The Institute for Genomic Research, 9712		
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	On Dec 8, 1999 this sequence version replaced g1:6102640.		
	Address all correspondence to:		
	Xiaoying Lin		
	The Institute for Genomic Research		
	9712 Medical Center Dr		
	Rockville, MD 20850, USA		
	E-mail: xlin@tigr.org		
	BAC clone FIM20 is from Arabidopsis chromosome I and is near the		
	molecular marker m1425.		
	The orientation of the sequence is from SP6 to T7 end of the BAC		
	clone.		
	Genes were identified by a combination of three methods: Gene		
	prediction programs including GRILL (available by anonymous ftp		
	from artuhr.spm.ornl.gov), GeneFinder (Phil Green, University Of		
	Washington), Genscan (Chris Burge,		
	http://www.cse.stanford.edu/~chris/GSNSCANv.html), and NetPlantGene		
	(http://www.obs.dtu.dk/netgene/cbsnetgene.html), searches of the		
	complete sequence against a peptide database and the Arabidopsis		
	EST database at TIGR (http://www.tigr.org/tdb/at/est.html).		
	Annotated genes are named to indicate the level of evidence for		
	their annotation. Genes with similarity to other proteins are named		
	after the database hits. Genes without significant peptide		
	similarity but with EST similarity are named as 'unknown' proteins.		
	Genes without protein or EST similarity, that are predicted by more		
	than two gene prediction programs over most of their length are		
	annotated as 'hypothetical'. proteins. Genes encoding tRNAs are		
	predicted by tRNAscan-SF (Sean Eddy		
	http://genome.wustl.edu/eddy/tRNAscan-SF/). Simple repeats are		
	identified by repeatmasker (Alan Smlt,		



genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

```
FEATURES
source
    /organism="Arabidopsis thaliana"
    /cultivar="Columbia"
    /db_xref="taxon:3702"
    /chromosome="I"
    /map="m1425"
    /clone="FIM20"
    138..269
/misc_feature
    /note="exon predicted by xgrail, quality
    marginal_shadowexon"
    complement(229..281)
/misc_feature
    /note="exon predicted by xgrail, quality excellent"
    complement(393..453)
/misc_feature
    /note="exon predicted by xgrail, quality excellent"
    complement(744..1286)
/misc_feature
    /note="exon predicted by xgrail, quality excellent"
    1271..1354
/misc_feature
    /note="exon predicted by xgrail, quality good_shadowexon"
    complement(1356..1475)
/misc_feature
    /note="exon predicted by xgrail, quality excellent"
    1378..1581
/misc_feature
    /note="exon predicted by xgrail, quality good_shadowexon"
    complement(1509..1530)
/misc_feature
    /note="exon predicted by xgrail, quality marginal"
    complement(1519<1655..1777,1908..2135,2258..2581,
    2679..2963,3052..>3429)
    /gene="FIM20.1"
    complement(<1655..>3429)
    /gene="FIM20.1"
    /note="similar to p58 protein kinase GB:AAB59449 [Homo
    sapiens]: contains Pfam profile: PF00069 Eukaryotic
    protein kinase domain"
    complement(1619<1655..1777,1908..2135,2258..2581,
    2679..2963,3052..>3429)
    /gene="FIM20.1"
    /codon_start=1
    /product="putative protein kinase"
    /protein_id="AAFI5905.1"
    /db_xref="GI:6539235"
    /translation="MGCVSKQVSVTPADISGVKDNENCSGSGRIYVEDPRPI
    LKTIIVSRSSKRSQKSGSELSSGASLSLSTRGNVSRTLEACVAGPAML
    SVNAGEILHGWVLRKEDAEKLEKIGQITSVNFAVETEGRIVALKVRDFEPE
    SYFMAKEILILRLNHPNLIKLEGLITKSCNQLVETEMHDLTGLSPDIKFT
    TPQIKCMKQILSGLDHCHSRGVMHDIKSGMLLSNEGILLVADGILANSNSGSK
    KRPILSRVYVILTRPELLIGATDYGASVDMVSGVFAELLGKPIRGRTVEOLH
    KIRKLGSPEDYWKSKLPHAMLFKPOQTYSCLEKTLKOLSEINILITLISIDP
    HKRGTAASALVSGYFKTKTPADPSLPIYPSKEIDTRHDEARVYSITFT"
    complement(4069..4131)
/misc_feature
    /note="exon predicted by xgrail, quality excellent"
    complement(join(4608..4760,4885..>4974))
    /gene="FIM20.2"
    complement(<4608..>4974)
    /gene="FIM20.2"
    /note="similar to dolichyl-phosphate mannosyltransferase
    polypeptide 2 GB:4503365 [Homo sapiens] (regulator of
    dolichol phosphate-mannose synthetase); EMBL J 1998 Sep
    1;17(17):4920-9)"
    complement(join(4608..4760,4885..4974))
    /gene="FIM20.2"
    /codon_start=1
    /product="putative dolichyl-phosphate mannosyltransferase
    polypeptide 2"
    /protein_id="AAFI5906.1"
    /db_xref="GI:6539236"
    /translation="MELADAVGLLSISLSIFYYFWIILPFVDSDFHRYFL
    PQYALIVPFAIGALISLISYFISGMVLMKSKRKA"
    complement(5111..5190)
    /note="exon predicted by xgrail, quality excellent"
    complement(<5342..>7603)
    /gene="FIM20.3"
    gene
        complement(<5342..>7603)
        /gene="FIM20.3"
        /note="contains Pfam profile: PF01348 Type II intron
        maturase"
        complement(5342..7603)
        /gene="FIM20.3"
        /codon_start=1
        /product="putative type II intron maturase"
        /protein_id="AAFI5907.1"
        /db_xref="GI:6539237"
        /translation="MASEKMGESLAGELIAYVESSSHYDDSKRSPRMEKRSIEL
        KKKRYKQCIKNGKPSDLKAYIRAPETLRDADCLIKNSVNSTENSGSVADSTAE
        ELSSGVFAVAKTESIVARDKTEVLVLPVAKLVQEARIVLEVVFSPFKISHS
        CSKSGGRASALKTINNNTSRSDMOCFTLSLNKKLDVSEFNLVMEKESDLSILL
        RSMFARVNLNLFEGGFPKGHLPOEGVLSRVNLMIYDRDFHEFYRISMREALGDS
        KRDSPGSKLMSFMRQAGBOGLSTEDVADLVAYCCRMEIYFSGPKVASD
        ISESLIGFRNSLHLDITDETPDPSCEKTSGLVGLVKNVRESPTVAVAKLKEK
        VRLAOKEMAMTIGTVRIQKKWGHGKVKYKESEITGGLADNSTLSQISCHAKME
        TDRWKILIRIMEDVYATSDRSBEVYLSKHVETPVPOELDATYRFQNRKAATYS
        STANLEALPOQSHDRPVFFGVAPATNIGKRLYRGLITAKGARSNSMLILD
        TAOIIDWVSGVLRWVWYEGCSNPFDEKALIDQIMSCIRTLAAKRTHEMEIKR
        LDLELSTPSAEDIEQELQHKRLDSPAFDRDEHLTVGLSNSGLCLSLATVSESPC
        NCFVIGCSMAAPVYTLHAMERORFPGWKTGFSVCISSINGRRIGICRKLDTYGC
        QISQAVDFGAWR"
        complement(7885..8010)
        /note="exon predicted by xgrail, quality marginal"
        8771..8821
        /note="exon predicted by xgrail, quality
        marginal_shadowexon"
        8935..9008
        /note="exon predicted by xgrail, quality
        marginal_shadowexon"
        9421..9528
        /rpt_family="POLY A"
        9761..9912
        /note="exon predicted by xgrail, quality marginal"
        10574..10662
        /note="exon predicted by xgrail, quality good"
        <10992..>14231
        /gene="FIM20.4"
        <10992..>14231
        /gene="FIM20.4"
        /codon_start=1
        /product="putative receptor protein kinase"
        /protein_id="AAFI5908.1"
        /db_xref="GI:6539238"
        /translation="MAATAVAGDSLSDSREVLTLSKYSLESRNPONGLTTEWMENO
        DVCVCGGILTCPPORSRYTGILNDSITSGFLFNFPAFLYTLDSRTTSGEIPD
        DISRCHNKLHNLSHNIEGELSLPGLSNLEVLDSLNRITGDIQSSFLCSNIYA
        NISTNFTGRIDIDFNGCRNKKYVDSNRSNGSGVWPGKRIYFSAVDNHSNTSA
        SMERKCTIQHMLDISGNAFGSGEPQVNSCNQNLAVLDMKRTKRTGIPATIGTSISK
        GILGNNTPSRDIPETLNLTLNVLFLDLSNRKTFGGDIOETGRFTQVKYVLHANSYV
        GGIINSNNIKLPLESRDLGYNFSQGLPTFISIOQLKPLILAIYNNFSGDIPQEVN
        MEGLOALDSEFKLNGSIPASFGKITSLMLMANNLSGEPREIENCSLMTFVYA
        NMLQSGREPELDTRMGNSNPSTFEVNRQNKIKIAGSGECLAMKRWIPAFPEFNVY
        AILTKRSKSLMDVILKGYGLFVCSASVYRTIKIAYLOLSGNKRSGEIPASISOM
        DRSLTHLGCFNEFEKGLPEIGLOLAFVNLNTRNPSGELIPQOELGNLCIONLDFN
        NFGSNPTSLNDLNLKFNISYNPETSGALIPYTGVAATLQKDSGLCNPLRPPSPFN
        OSGNTRKISNOYLCNRRPTIILWISLALAIACILAVISGYLWYAKASRAEIDL
        LGSKTRBDMKISSGSSPWSGKITVIRLKSSTFYADIKATISNSEKVRGQGY
        GTVYRGVLPDGRVAVKRLQRBGTEAEKEFAEMEVISANAGDMAPNLVRLYGMCL
        DSEKILVHEYVGGSLDELITDRKFLQMKRRDIIADVADLGVFLHHECPSTIARD
        VASNNLIDKHGNARVTFGLARLNVGDSHVSIVAGTIGYVAPEYQWQATFTGCD
```

```
gene
    complement(<5342..>7603)
    /gene="FIM20.3"
    /note="contains Pfam profile: PF01348 Type II intron
    maturase"
    complement(5342..7603)
    /gene="FIM20.3"
    /codon_start=1
    /product="putative type II intron maturase"
    /protein_id="AAFI5907.1"
    /db_xref="GI:6539237"
    /translation="MASEKMGESLAGELIAYVESSSHYDDSKRSPRMEKRSIEL
    KKKRYKQCIKNGKPSDLKAYIRAPETLRDADCLIKNSVNSTENSGSVADSTAE
    ELSSGVFAVAKTESIVARDKTEVLVLPVAKLVQEARIVLEVVFSPFKISHS
    CSKSGGRASALKTINNNTSRSDMOCFTLSLNKKLDVSEFNLVMEKESDLSILL
    RSMFARVNLNLFEGGFPKGHLPOEGVLSRVNLMIYDRDFHEFYRISMREALGDS
    KRDSPGSKLMSFMRQAGBOGLSTEDVADLVAYCCRMEIYFSGPKVASD
    ISESLIGFRNSLHLDITDETPDPSCEKTSGLVGLVKNVRESPTVAVAKLKEK
    VRLAOKEMAMTIGTVRIQKKWGHGKVKYKESEITGGLADNSTLSQISCHAKME
    TDRWKILIRIMEDVYATSDRSBEVYLSKHVETPVPOELDATYRFQNRKAATYS
    STANLEALPOQSHDRPVFFGVAPATNIGKRLYRGLITAKGARSNSMLILD
    TAOIIDWVSGVLRWVWYEGCSNPFDEKALIDQIMSCIRTLAAKRTHEMEIKR
    LDLELSTPSAEDIEQELQHKRLDSPAFDRDEHLTVGLSNSGLCLSLATVSESPC
    NCFVIGCSMAAPVYTLHAMERORFPGWKTGFSVCISSINGRRIGICRKLDTYGC
    QISQAVDFGAWR"
    complement(7885..8010)
    /note="exon predicted by xgrail, quality marginal"
    8771..8821
    /note="exon predicted by xgrail, quality
    marginal_shadowexon"
    8935..9008
    /note="exon predicted by xgrail, quality
    marginal_shadowexon"
    9421..9528
    /rpt_family="POLY A"
    9761..9912
    /note="exon predicted by xgrail, quality marginal"
    10574..10662
    /note="exon predicted by xgrail, quality good"
    <10992..>14231
    /gene="FIM20.4"
    <10992..>14231
    /gene="FIM20.4"
    /codon_start=1
    /product="putative receptor protein kinase"
    /protein_id="AAFI5908.1"
    /db_xref="GI:6539238"
    /translation="MAATAVAGDSLSDSREVLTLSKYSLESRNPONGLTTEWMENO
    DVCVCGGILTCPPORSRYTGILNDSITSGFLFNFPAFLYTLDSRTTSGEIPD
    DISRCHNKLHNLSHNIEGELSLPGLSNLEVLDSLNRITGDIQSSFLCSNIYA
    NISTNFTGRIDIDFNGCRNKKYVDSNRSNGSGVWPGKRIYFSAVDNHSNTSA
    SMERKCTIQHMLDISGNAFGSGEPQVNSCNQNLAVLDMKRTKRTGIPATIGTSISK
    GILGNNTPSRDIPETLNLTLNVLFLDLSNRKTFGGDIOETGRFTQVKYVLHANSYV
    GGIINSNNIKLPLESRDLGYNFSQGLPTFISIOQLKPLILAIYNNFSGDIPQEVN
    MEGLOALDSEFKLNGSIPASFGKITSLMLMANNLSGEPREIENCSLMTFVYA
    NMLQSGREPELDTRMGNSNPSTFEVNRQNKIKIAGSGECLAMKRWIPAFPEFNVY
    AILTKRSKSLMDVILKGYGLFVCSASVYRTIKIAYLOLSGNKRSGEIPASISOM
    DRSLTHLGCFNEFEKGLPEIGLOLAFVNLNTRNPSGELIPQOELGNLCIONLDFN
    NFGSNPTSLNDLNLKFNISYNPETSGALIPYTGVAATLQKDSGLCNPLRPPSPFN
    OSGNTRKISNOYLCNRRPTIILWISLALAIACILAVISGYLWYAKASRAEIDL
    LGSKTRBDMKISSGSSPWSGKITVIRLKSSTFYADIKATISNSEKVRGQGY
    GTVYRGVLPDGRVAVKRLQRBGTEAEKEFAEMEVISANAGDMAPNLVRLYGMCL
    DSEKILVHEYVGGSLDELITDRKFLQMKRRDIIADVADLGVFLHHECPSTIARD
    VASNNLIDKHGNARVTFGLARLNVGDSHVSIVAGTIGYVAPEYQWQATFTGCD
```

Query Match 10.0%; Score 79.2; DB 8; Length 134402;  
Best Local Similarity 52.8%; Pred. No. 6 4e-12;  
Matches 171; Conservative 0; Mismatches 153; Indels 0; Gaps 0;



OY	231	tgagttctgggaaattcaaacccctcttgacactcttgccctggagcttgggaaacttgaaac	290
Db	12043	TAGATTGGGCTACAAACATTTCTGGGACAGTACCTACCTAGAAATTTCTCAGATACGAA	12102
OY	291	attacagatctagaagctctcaaaaaaacatccaagaagacatactctccgaacttg	350
Db	12103	GTTTGAAGTTCTTGATTTCTGCTTATTAATACCTCAGTGGGAAATTCACAGAGATTAG	12162
OY	351	gaactcgaagaatccatccatcagacttgagatctgtaacaacaacatctacagagatgac	410
Db	12153	GGAACATGCGGGGGCTTCAAGCAGCTTGATCTCTCTTTAAACAGCTGACCGGTTGATAC	12222
OY	411	ccactctcttgggaaaaattgaaagctctcggfctctttaagcgcttaatgaaacccgatgaa	470
Db	12223	CAGTTTCAATTTGGGAAATTTGACCTCTCTTTGTGGCTAATGCTTGCAAAACAACTCTCAT	12282
OY	471	cgaggccaactccctagaagacactcaactgcatccacagcccttaaaagtgttgatgctcaca	530
Db	12283	CAGGAGAAATTCCTCGAGGAGATTGGCAACATGCACAAAGCCTTTGTGGTTTAAGTGGCAA	12342
OY	531	gcactgattctggtgaaacaacac	554
Db	12343	ACAAACAAGCTCTCTGGTAGAATTC	12366

Search completed: June 24, 2000, 00:21:39  
Job time: 46254 sec







```

Db      |||||||
361 GAATCATCATGAGTGGATCTGTACAACAACATCTTACAGGAGATAGTCCACCTCTTT 420
QY      |||||||
421 gggaaatgaagtctctgtctttaaagcctaagcaacggttaagggccaat 480
Db      |||||||
421 GGGAAATGAAGTCTGTGCTTTTACGGCTTAATGACACCGATGACGGGCCAAT 480
QY      |||||||
481 ccttaagcagctacatgcaatcccaagccttaagttgtatgctcaagcaatgatt 540
Db      |||||||
481 CCGTAGAGCACTACAGCAATCCCAAGCCTTAAGTTGATGATGCTCAAGCAATGATT 540
QY      |||||||
541 gtttgaagaacatcccaagaagacctttgtctacattcccttaacaatttgagaa 600
Db      |||||||
541 GTTGGAAACATCCCAACCAACGACCTTTGTCTCATCTCTTACAGAACTTTGAGAA 600
QY      |||||||
601 caaccgaggttgaggaagcaggaattactcgtcttgcaagctaacacatactgac 660
Db      |||||||
601 CAACCCAGGTTGAGGAGGACCGGAATTACTCGGCTTTCAGAGTACGACTACTGCAC 660
QY      |||||||
661 ctgaataaattggcaaaaacctgaataatgagaatttgggggtgaacctgttaagaacctt 720
Db      |||||||
661 CTGAAAAAATTTGGCAAAACCTGAAATGAGAAATGGGGGTGACCTTGAAGAACACTT 720
QY      |||||||
721 caccacttcaaatatcacatcactactatgtaataagatatataatgtagtccaaaaa 780
Db      |||||||
721 CACCACCTTTATCAATATTCACATCTACTATGTAATAGTATATATATATATGATCCAAAAA 780
QY      |||||||
781 aaaaaaaa 788
Db      |||||||
781 AAAAAAAA 788

RESULT 2
V06590 ID V06590 standard; cDNA to mRNA; 1063 BP.
AC V06590;
AT 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS 106..762
FT FT /tag- a
FT FT /note- "shows high homology to SERK"
PN W09743427-A1.
PD 20-NOV-1997.
PE 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
P-PSDB: W47022.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 86-88; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;

Query Match 97.5%; Score 769.6; DB 1; Length 1063;
Best Local Similarity 98.9%; Pred. No. 9.8e-223;
Matches 775; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Db      |||||||
223 AGTTAACAGATCCGGACATGTTCTCCAGGTGGATCCAACTTGTATACCTTGT 282
QY      |||||||
185 acctgttccatgtacactgttaaccaagaacccggtcactcgtgtgatttgggaaat 244
Db      |||||||
283 ACCGTGTTCCATGTACACCTTAAACCAAGAACCCGGGTACCTGTTGATTTGGGAAAT 342
QY      |||||||
245 tcaaacctctcgtgaacatcttcgacctgagcttgggaagcttgaacatttaagttctca 304
Db      |||||||
343 TCAAACTCTCTGTGACATCTTGGCCCTGAGACTTGGGAAGCTTGAACATTACAGTATCA 402
QY      |||||||
305 gagcttacaacaaacacatcccaaggaacatataccttcggaacttggaaatcgtgaagt 364
Db      |||||||
403 GAGCTTACAAAACCAATCCCAAGAACATACCTTCGAACTTGGAAATCGAAGAAAT 462
QY      |||||||
365 ctacacagcttgatctgtacaaacaaatcttaacaggatagttccactcttggga 424
Db      |||||||
463 CTCATCAGCTTGATCTGTACAAACAAATCTTACAGGGATAGTCCACTTCTTTGGGA 522
QY      |||||||
425 aaattgaagctctgtcttctttaaagcctaataagcaacagcattgagcgggccaatccct 484
Db      |||||||
523 AAATTGAAGTCTGTGCTTTTACGGCTTAATGACACCGAATGACGGGCCAAATCCCT 582
QY      |||||||
485 agagcactcactgcaatcccaagccttaaaagtgttgaatgtctcacaagcaatgatttgt 544
Db      |||||||
583 AGAGCACTCACTGCAATCCCAAGCCTTAAAGTTGTTGATGTCTCAAGCAATGATTGTGT 642
QY      |||||||
545 ggaacaatcccaacaaacaggaacctttgtctacatctccttaagaaccttggaaacaac 604
Db      |||||||
643 GGAACAAATCCCAACAAACGACCTTTGTGCTCACATTCCTTTACAGAACTTGAACAAAC 702
QY      |||||||
605 ccgaggttgaggaagcaggaattactcgtcttgcaagctacagacactaactcacctga 664
Db      |||||||
703 CCGAGTTGAGGAGACCGGAATTACTCGGCTTTCAGACTACGACATACCTACCTGA 762
QY      |||||||
665 aaaaattggcaaaacctgaataatgagaatttgggggtgaccttgaagaacacctacc 724
Db      |||||||
763 AAAAATTTGGCAAAACCTGAAATGAGAAATGGGGGTGACCTTGTGAACACTTCCACC 822
QY      |||||||
725 acttataaataatcacatcactactatgtaataagtatataatgtagtccaaaaaaa 784
Db      |||||||
823 ACTTATCAAAATTCACATCTACTATGTAATAGTATATATATATATATGACTCCAAAAAAA 882
QY      |||||||
785 aaaa 788
Db      |||||||
883 TGA 886

RESULT 3
V06587 ID V06587 standard; cDNA to mRNA; 981 BP.
AC V06587;
AT 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS 104..760
FT FT /tag- a
FT FT /note- "shows high homology to SERK"
PN W09743427-A1.
PD 20-NOV-1997.
PE 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
P-PSDB: W47019.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 75-77; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.

```



Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 94.0%; Score 742; DB 1; Length 981;  
Best Local Similarity 96.8%; Pred. No. 26-214;  
Matches 757; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

5 cccacgctccggaacatacgtgagagctctcgagcttcgttaacttaactta 64  
101 CACATGCGCTCGAACAATCGGTGGAGCTTCGCAAGCTTCGTAACCTTA 160  
65 gcttgaattacactggtcggaagcaactcggaaggaagctcttaacttcg 124  
161 GCTTGATTCACCTGGTGGAGCAACTCGAGAGGAGATCTTCACTCTTGG 220  
125 agtttaacagatccggaacatgcttcacagagctggatccaactctgt 184  
221 AGTTGACAGATCCAGACCATGCTCCAGAGCTGGGATCCAACTCTGTAT 280  
185 accgtgtccatgctacactgtaacaaagacacggtactcgtgtgattgg 244  
281 ACCTGTTCCATGTCACCTGTAAACCAAGACACCGGCTCCTGTGATTTGG 340  
245 tcaaacctctctgacatacttcgcgcctgagcttgggaagcttgaacat 304  
341 TCAAACTCTCTGGACATCTTGGCCCTGAGCTTGGGAGCTTGAACATTAC 400  
305 gagctctacaacaaacacatccaaaggaactataccttcggaacttga 364  
401 GAGCTTCAACAAAACCAATCCAAAGGAATATCTTCCAACTTGGAAATCG 460  
365 ctcatcgactgagatctgtacaacaacatctcacagaggaagtccact 424  
461 CTCTACAGCTTGATCTGTACAAACAATCTTACAGGATTAATCTCCACT 520  
425 aaattgaagctctgctctttaaagcttaagaaacccgattgacggagg 484  
521 AAATTGAAGTCTCGGTCTTTTACGGCTTAATGACAAACGATGACCGGT 580  
485 agagcaactcactgcaatcccaagcttaaaagtgtgatttccaaagca 544  
581 AGAGCACTCAACGGCAATCCAAAGCTTAAAGTTGTACCTTCAAGCAAT 640  
545 ggaacaatcccaacaaacaggaacttgcacatctccttacaagaactt 604  
641 GGAACAATCCCAAAACGGAACCTTGTACATCTTACAGAACTTGAACAAC 700  
605 ccgaggtctggagggacccggaattactcgtgtcttgcgaagctaac 664  
701 CCGAGATTGGAGGGACCGGAATTAATGCTTGCACAGCACTAACTCACCT 760  
665 aaaaattgcaaaaccttgaaaatgaagaatttgggggtgaccttgaaga 724  
761 AAACAATCGCAAAACCTGAAAATGAAGATTTGGGGGTGACCTTTGAAG 820  
725 acttataaataatcacactacatgataaagatatatagtgtccaaaaa 784  
821 ACTTATCAAAATATCATCTATTATTAATAGTATATATGATAAAACAAA 880  
785 aa 786  
881 AA 882

RESULT 4  
ID V06586  
V06586 standard; cDNA to mRNA; 1106 BP.  
AC V06586;  
DT 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK IRR homologous EST clone.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KW plant breeding; leucine-rich repeat; sr.  
OS Arabidopsis thaliana.

Key Location/Qualifiers  
FT 142..798  
FT /tag=a  
CDS /note="shows high homology to SERK"

MO9743427-A1.  
PN 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PI 14-MAY-1996; GB-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt BDL, Van Holst GJ.  
DR WPI: 98-086529/08.  
DR P-PSDB: M47018.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 28; Pages 71-73; 123pp; English.  
CC The sequence is that of an EST clone showing high homology to  
CC SERK IRR (leucine-rich repeat) sequences.  
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;

Query Match 90.5%; Score 714.2; DB 1; Length 1106;  
Best Local Similarity 96.2%; Pred. No. 5.2e-206;  
Matches 753; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

5 cccacgctccggaacatacgtgagagctctcgagcttcgttaacttaact 64  
139 CACATGCGCTCGAACAATCGGTGGAGCTTCGCAAGCTTGAACCTTA 198  
65 gcttgaattacactggtcggaagcaactccggaaggaagatgctcttaac 124  
199 GCTTGATTCACCTGGTGGAGCAACTCGAGAGGATGCTTACGCTTGGCGG 258  
125 agtttaacagatccggaacatgcttcacagagctggatccaactctgt 184  
259 AGTTGACAGATCCAGACCATGCTTCCAGAGCTGGGATCCAACTTGTAT 318  
185 accgtgtccatgctacactgtaacaaagacacggtactcgtgtgattgg 244  
319 ACCGTTTCATGTCACCTGTAAACCAAGACACCGGCTCCTGTGATTTGG 378  
245 tcaaacctctctgacatacttcgcgcctgagcttgggaagcttgaacat 304  
379 TCAAACTCTCTGGACATCTTGGCCCTGAGCTTGGGAGACTTGAACATT 438  
305 gagctctacaacaaacacatccaaaggaactataccttcggaacttga 364  
439 GAGCTTCAACAAAACCAATCCAAAGGAATATCTTCCAACTTGGAAATCG 498  
365 ctcatcgactgagatctgtacaacaacatcttaagggatagttccact 424  
499 CTCTACAGCTTGATCTGTACAAACAATCTTACAGGATTAATGCTTGGGA 558  
425 aaattgaagctctgctctttaaagcttaagaaacccgattgacggagg 484  
559 AAATTGAAGTCTCGGTCTTTTACGGCTTAATGACAAACGATGACCGGT 617  
485 agagcaactcactgcaatcccaagctttaaagtgtgtgatttccaaag 543  
618 AGAGCACTCAACGGCAATCCCAAGCCTTTAAAGTTGTGACGCTCAAGCA 677  
544 tgaacaatcccaacaaacggaacttctgctacatctccttacaagaact 603  
678 TGGGACAATCCCAACAAGGACCTTGTGTCATCTTACAGAACTTGAAGACA 737  
604 ccggaagtggagggacccggaattactcgtgtcttgcgaagctaacact 663  
738 CCGGAGATTGGAGGGACCGGAATTAATGCTTGCACAGCACTAACTCACCT 797  
664 aaaaattgcaaaaccttgaaaatgaagaatttgggggtgaccttgaaga 723  
798 AAACAATCGCAAAACCTGAAAATGAAGATTTGGGGGTGACCTTGAAGCA 857  
724 caattataaataatcacactacatgataaagatatatagtgtgtgtgt 783



D	b	858	CACCTTATCAAAATTCACACTATATTATGTAATAAGTATATATGTAGTAAAAA	917
Oy		784	aaa 786 	
D	b	918	AAA 920	
<hr/>				
RESULT      5				
ID		V06589	standard; cDNA to mRNA; 894 BP.	
AC		V06589;		
DY		03-AUG-1998	(first entry)	
DE		Arabidopsis thaliana SERK LRR homologous EST clone.		
KW		receptor kinases; apomixis; apomictic seeds; production; embryos;		
KW		plant breeding; leucine-rich repeat; ss.		
OS		Arabidopsis thaliana.		
FH		key	Location/Qualifiers	
FT		1..678		
FT		CDS		
FN		MO9743427-AI.	/note= "shows high homology to SERK"	
PD		20-NOV-1997.		
PE		13-MAY-1997; E02443.		
PR		14-MAR-1996; GB-010044.		
PA		(NOVS ) NOVARTIS AG.		
PI		De Vries SC, Hecht vFG, Schmidt EDL, Van Holst GJ;		
DR		WPI: 98-086529/08.		
P		P-PSDB: W47021.		
PT		Production of apomictic seeds - useful in plant breeding		
PS		Claim 28; Pages 83-84; 123pp: English.		
CC		The sequence is that of an EST clone showing high homology to		
CC		SERK LRR (leucine-rich repeat) sequences.		
SO		Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;		
<hr/>				
Query Match                  45.8%; Score 361; DB 1; Length 894; Best Local Similarity 72.8%; Pred. No. 1,8e+99;				
Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2				
OY		91	cctccgaagagaatgctccttaacgccttcgccggagttaaacaagaccgacatgttct 150 	
Db		15	CTCCGAAGGGGAGCGCCTCTTCACCGCCTTGCCGGAGCTTATCAGATCCAGACAATGTTGT 74 	
OY		151	cctaagcttggaatccaactctgtttaacctgttaacctgttcaactgttaacca 210 	
Db		75	TCAAGATGGGATCCAACCTGTTATATCTGTGACTGGTTTCATGCTCATCTGTATACA 134 	
OY		211	agacaacogcgctcaactcgttgttgatttgaggaaattcaaacctctctgpaactcttgcc 270 	
Db		135	ACACACATCAAGTACTACTGCTGTGATTGGGGAATTCAAACTATCTGGACATCTAGTACC 194 	
OY		271	tgaagcttggaagacttgaaacttaacagtatct----- 304 	
Db		195	TGAAGTGGGGAAGCTTGAMCATTTACAAATATCTGTATGATGAAATCATCACTCTTTGCCCTT 254 	
OY		304	----- 304	
Db		255	TGATTATCTGAAAACATTTCATTATCACGTACACACATATAACATTTCCTTTGAGTCATA 314 	
OY		304	---agagctctcaaaaaaacaataccaaggaactatactctccgaacttggaaacttgaa 360 	
Db		315	TAGTGAATCTCAAAAAACGAGAATTCAGAGAACATATACCTTCTGAGCTTGGAAATCTGAA 374 	
OY		361	gaatctcatagcttgtagcttgfaaacaacaacttaagggaagttacatctctt 420 	
Db		375	GAGTCTAATCTAGTTGGATCTGTACACACACAACTCACCGGGAAAAATCCATCTTCTTT 434 	
OY		421	gggaatatggaagctctcgtgtcttttacggcttaatgaaacacgataggagggccaat 480 	
Db		435	GGGAAATTTGAAGTCACTGTTTTTTTGGGGCTTAACGAAAAACGATTTGACCGGATCTAT 494 	
OY		481	ccttagagagctcaactcgaatcccgaagccttaagttgttgattgtctcaagaatgattt 540 	

[illegible]



Db	330	GTCTTCAGACGCTGGGATTCACCGCTAGTGAATCCTTGACATGGTTCATGTCACCTTGC	389
Qy	206	aaccaagaacaaccgcgcactcgtgtggaatttgggaattcaaccctcttgcacatct	265
Db	390	AACAAAGAAACAGAGTCTATAGAGTGTGATTTGGGAATCAGAGAGTATCTCGCCATTTA	449
Qy	266	gcgcctgcagcttgggaagcttgcacatttcagtatctagaacttaaaaaaacaatc	325
Db	450	GTTCACGACCTGGGTGCTCTAAAGAAATTTGCAGTATTGTGACCTTTAAGTAAACATA	509
Qy	326	caaggaacataccttcgcgaacttggaaatctgaagatcctatagcttgcgtctac	385
Db	510	ACTGGCCCATTCCTAGTAATCTTGGAAATCTACAAACTTGTGAAGTTTGGATCTTTAC	569
Qy	386	aacaacaatcttacagggatagttcccaactctcttggaaaaattgaagtcctgtcttc	445
Db	570	TTAAACAGGTTCTCCGGGTCATTTCCGGGATCTATTGGGAAAGCTTTCAAAGCTGAGTTT	629
Qy	446	ttaaggcttaatgacaaccgattgacgggggccaatcccttaagacacatcgtgcaatcca	505
Db	630	CTCCGGCTTAACAAACAAACAGTCACTACGGGGCAATTCATGTCACATGACCAATATTACT	689
Qy	506	agaccttaagttgttgatgctcaagcaatgatttgttggaaacaatcccaacaacogga	565
Db	690	ACCCTTCAGAGTGTGAGTCATCAAAATTAACAGACTCTGCTGTTCAGTTCTCGAACATGGC	749
Qy	566	cccttgcctcacatcccttcaagaagaaacttggagaagacac	604
Db	750	TCCCTTCACCTTTACACCCATCATGATTTTGGCTATTAAC	788

RESULT	7
ID	V06571
AC	V06571; standard; CDNA; 1814 BP.
DT	03-AUG-1998 (first entry)
DE	Daucus carota SERK gene.
KM	receptor kinase; apomixis; apomictic; seeds; production; embryos
KW	plant breeding; ss.
OS	Daucus carota.
PH	key
FT	Location/Qualifiers
FT	94..1755
FT	/*tag= a
PN	/product= SERK protein
PD	W09743427-A1.
PF	13-MAY-1997.
PR	14-MAY-1996; E02443.
PA	(NOVS ) NOVARTIS AG.
PI	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;
DR	WPI: 98-086529/08.
DR	P-PSDB: W47013.
PT	Production of apomictic seeds - useful in plant breeding
PS	Claim 21; Pages 47-51; 123pp; English.
CC	The sequence is that encoding SERK, a putative receptor kinase.
CC	It may be used as part of a method of producing apomictic seeds
CC	comprising: (a) transforming plant material with a nucleotide
CC	sequence encoding a protein which in active form in a cell or
CC	cell membrane renders the cell embryogenic; (b) regenerating
CC	the transformed material into plants or carpel-containing
CC	plant parts; and (c) expressing the sequence in the vicinity
CC	of the embryo sac. The apomictic seeds and embryos thus produced
CC	can be developed into plant progeny. This is useful in plant
CC	breeding programs. Controllable and reproducible apomixis providd
CC	many advantages in plant improvement and cultivar development in
CC	the case that sexual plants are available as crosses with the
CC	apomictic plant. Apomixis provides for true-breeding, seed
CC	propagated hybrids and could shorten and simplify the breeding
CC	process so that selfing and progeny testing to produce and/or
CC	stabilise a desirable gene combination could be eliminated.
CC	Apomixis allows plant breeders to develop cultivars with
CC	specific stable traits for such characteristics as height,

[illegible]

QY	265	tggcgcttgagcttgggaagcttgaacatttaacgatatccagagctctacaacaaacacat	324
Db	138	TGATGCTTACCTTGACAAATATGGGGTCTTATGACATTTGGAGCTTTACACCAATTAACAT	197
QY	325	ccaaggaactacacctccgaacttggaaactctgaagaactcaccacagcttgcgtla	384
Db	198	AAATGGACCAATTCCTAGTAGTCTTGGGAAATGCAAAATTTGGAGCTTGAGCACTATA	257
QY	385	caacaaacaactctaaaggatagttcccaactctcttgggaaatttgaagtctctgctc	444
Db	258	CATGAATACCTTCTTGACACTATACCGGACACATTTGGAAAGCTTAAAGGCTTAAGATT	317
QY	445	tttaaggcttaatgacaacgcataggcggggccaatccctagagacatcgaatccg	504
Db	318	CTTGCGCTCTAACAAACAAGCCTCTCTGGTCCAAATTCCAATGTACATGACTAATATTAC	377
QY	505	aagcccttaaaagtgttgatgtgtctcaagcaatgattgtgtggaacaatccacaacacg	564
Db	378	AACTCTTCANAGTCGTGAAATTATCAACAATCCGGCTATCAGGACCAAGTACCGGAATATGG	437
QY	565	acactttgctcacatccctcttaacgaacatttgaaacaaacccgaggttggaggacgga	624
Db	438	CTCATTTTCTTTGTTTACACCTATCAGTTTGGCAATTAATTTGAATTTATGTGGGCCCGT	497
QY	625	A 625	
Db	498	A 498	

SEQUENCE	3176 Bp.
SD	927 A; 654 C; 659 G; 936 T;
CC	plant. The MRP's cDNA or genomic DNA can be used to transform a
CC	plant to increase its MRP expression, and therefore control the
CC	form (particularly stem length) of the plant.
CC	
CC	acts on a plant at a specific site for a specific period, and can
CC	a plant with, e.g. short stems or altered inflorescence. The MRP
CC	morphogenesis regulatory protein (MRP), which can be used to
CC	The present sequence encodes an Arabidopsis thaliana plant
PS	Claim 1; Pages 6-10; 17pp; Japanese.
PT	yield plants with short stems or altered inflorescence
FT	DNA encoding plant morphogenesis regulatory protein - useful to
DR	p-RSDS: W13408.
PA	(CHIR-) ZH CHIKUY KANKYO SANGYO GIJITSU KENKYU.
PR	24-AUG-1995; JP-216187.
PD	04-MAR-1997.
PN	J09056382-A.
FT	/tag-a
FT	/note-"Plant morphogenesis regulatory protein"
FH	Key location/Qualifiers
FF	cds 51..2981
ID	T62124 standard; cDNA to mRNA; 3176 Bp.
AC	T62124; 10-0UN-1997 (first entry)
DE	Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
KW	Plant; morphogenesis; regulation; short; stem; alteration;
KV	inflorescence; extraneous; gene; expression; transformation;
OS	increase; control; form; length; ds.
OS	Arabidopsis thaliana.



[illegible]

RESULT 9

ID X23526 X23526 standard; DNA; 5940 BP.

AC X23526;

DT 17-JUN-1999 (first entry)

DE O. longistaminata xaz1 gene family member A2 DNA.

KW xaz1; receptor kinase-like protein; multigene family; RRK; rice;  
plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.  
Oryza longistaminata.

OS Oryza longistaminata.

PN WO9909151-A2.

PD 25-FEB-1999.

PI 17-JUL-1998; U14841.

PR 13-AUG-1997; US-910386.

PA (REGC ) UNIV CALIFORNIA.

PI Hulbert SH, Richter T, Ronald PC, Song W, Szabo V,  
Mang G;  
WR: 99-204431/17.

DR New RRK polynucleotides and nucleic acid constructs - used for  
generating transgenic plants resistant to Xanthomonas

PT Claim 1; Page 52-53; 67pp; English.

PS This invention describes a method for conferring disease resistance in  
plants. The invention describes the use of novel genes and proteins  
belonging to the Oryza longistaminata and Oryza sativa receptor  
kinase-like protein (RRK) xaz1 multigene family. Such genes from  
cassava, maize and tomato are also described. The genes and proteins can  
be used for enhancing resistance to Xanthomonas in a plant, preferably in  
rice or tomato.

SO Sequence 5940 BP; 1570 A; 1200 C; 1188 G; 1982 T;

Query Match	9.58;	Score 75;	DB 1;	Length 5940;
Best Local Similarity	52.48;	Pred. NO. 8.8e-13;		
Matches 165;	Conservative	0;	Mismatches 150;	Indels 0;
			Gaps	0;

QY 245 tcaaacccctcttgacatcttgagccctgagcttgaggaaattgaacatttaacagatcta 304  
Db 2406 TCCGCGCTGTCGGGAGTCAATCTCCCTCCCTGGCACTGTCTCTCTCAGAGACGCT 24655  
QY 305 gagcctctcaaaaacaacatccaaaggaactataccttcggaacttggaaatctgaagaat 364  
Db 2466 GACCTCAGCAGCAACCAACCTGTCCGGCAGATGACCCAGAACTCAGCAGTCTCAGCAGG 25252  
QY 365 cctatcagcttggaatctggtacacacaacatcttaacaggaatagttcccaactcttggga 424

Accession	Sequence	Position
Db	CTCCAAACAATGGTACTGAAATTTCACACACGCTATGGGGAGATTTCCAGCTGCTTTGGG	2585
Db	2526	
QY	425 aaattgaagctctctgctgctctttaaeggttaatgacacgattgacggggccaatccct	484
Db	2586 AATTTAACCAAGTCTTCGGTCTTTGAGCTGACTAACAAATACCTCTCGGAGCAATCCCT	2645
QY	485 agagcattcaactcgtcatcccaacgacctaaagttgtttgatgctcgaagcaatgattgtgt	544
Db	2646 TCACATCTGGGGCAAACTCACACAGCCCTCACTGTGATCTTGCACTGGCTGAATAATATGCTGTCT	2705
QY	545 gnaacaaatcccaaca 559	
Db	2706 AGTTCATCTCCCTCA 2720	

RESULT 10

ID T06307

AC T06307; standard; cDNA; 3573 BP.

DT 14-APR-1996 (first entry)

DE Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.

KW Pathogen resistant; Cf-2.2; tomato; C.fulvum; AVR 4; AVR 9; fungal; leaf mould; variegation; ss.

OS Lycopersicon esculentum.

PN M0531564-A2.

PD 23-NOV-1995.

PE 11-MAY-1995; G01075.

PR 11-MAY-1994; GB-009394.

PR 23-DEC-1994; WO-G02812.

PR 31-MAR-1995; GB-006558.

PR 07-APR-1995; GB-007232.

PA (GATS-) GATSBY CHARITABLE FOUND.

P1 Hammond-Kosack KE, Jones DA, Jones JDS;

P2 WPI; 96-010849/01.

P3 P-PSDD; R85599.

P4 Increasing plant pathogen resistance by induction of variegation -

P5 may lead to acquired resistance to a broad range of pathogens.

PS Claim 9; Page 85-87; 131pp; English.

CC T06307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone.

CC which may be used to impart a broad range of pathogen resistance, by

CC induction of variegation, to transgenic plants (or parts or propagules

CC of plants) containing such constructs. Cf-2.2 imparts resistance to the

CC disease caused by the leaf mould fungal pathogen Cladosporium fulvum.

CC C.fulvum contains avirulence (Avr) genes that confer recognition by

CC plants containing Cf-genes, leading to the activation of host defence

CC mechanisms to attack the disease.

SQ Sequence 3573 BP; 1032 A; 654 C; 664 G; 1223 T;

Query Match	9.0%;	Score 71.4;	DB 1;	Length 3573;
Best Local Similarity	50.4%;	Pred. No. 8.5e-12;		
Matches 174;	conservative	0;	Mismatches 171;	Indels 0;
				Gaps 0.

[illegible]







CC process so that selfing and progeny testing to produce and/or  
 CC stabilise a desirable gene combination could be eliminated.  
 CC Apomixis allows plant breeders to develop cultivars with  
 CC specific stable traits for such characteristics as height,  
 CC seed and forage quality and maturity.  
 SQ Sequence 4081 Bp; 1120 A; 770 C; 785 G; 1406 T;

Query Match 9.0%; Score 70.8; DB 1; Length 4081;  
 Best Local Similarity 69.6%; Pred. No. 1.4e-11;  
 Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 97 aggaagatgcctcttaacgctcttcgcggaggttaacagatccggaccatgtctccagag 156  
 DB 1794 AGGAGATGCTTTGATCTTGAGGGTTACTGATGATCCAAACATGTCTTGCAAGAG 1853  
 QY 157 ctggagaccacactctgttaacctctgttaccctgttccatgttaactgttaaccaagaca 216  
 DB 1854 CTGGGATCCATACGCTAGTGAATCTTGCACATGTTCATGTCATGCAACACAGAA 1913  
 QY 217 ccgcgtcactcgtgtga 234  
 DB 1914 CAGTGTCTATAAGAGTGTGA 1931

## RESULT 13

ID V06570 standard; DNA; 6695 BP.

AC V06570;  
 DT 03-AUG-1998 (first entry)

DE Daucus carota SERK gene.

KM Receptor kinase; apomixis; apomictic; seeds; production; embryos;  
 KM plant breeding; ds.

OS Daucus carota.

FR Key Location/Qualifiers

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

WO9743427-A1.  
 PD 20-NOV-1997.  
 PR 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVUS) NOVARTIS AG.  
 DE Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 PI WPI: 98-086529/08.  
 DR P-PSDB; W47013.

PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 21: Pages 40-46; 123pp; English.

CC The sequence is that encoding SERK, a putative receptor kinase.

CC It may be used as part of a method of producing apomictic seeds

CC comprising: (a) transforming plant material with a nucleotide

CC sequence encoding a protein which in active form in a cell or

CC cell membrane renders the cell embryogenic; (b) regenerating

CC the transformed material into plants or carpel-containing

CC plant parts; and (c) expressing the sequence in the vicinity

CC of the embryo sac. The apomictic seeds and embryos thus produced

CC can be developed into plant progeny. This is useful in plant

CC breeding programs. Controllable and reproducible apomixis provides

CC many advantages in plant improvement and cultivar development in

CC the case that sexual plants are available as crosses with the

CC apomictic plant. Apomixis provides for true-breeding, seed

CC propagated hybrids and could shorten and simplify the breeding

CC process so that selfing and progeny testing to produce and/or

CC stabilise a desirable gene combination could be eliminated.

CC Apomixis allows plant breeders to develop cultivars with

CC specific stable traits for such characteristics as height,  
 CC seed and forage quality and maturity.

SQ Sequence 6695 Bp; 1844 A; 1182 C; 1243 G; 2422 T;

Query Match 9.0%; Score 70.8; DB 1; Length 6695;  
 Best Local Similarity 69.6%; Pred. No. 1.7e-11;  
 Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 97 aggaagatgcctcttaacgctcttcgcggaggttaacagatccggaccatgtctccagag 156  
 DB 956 AGGAGATGCTTTGATCTTGAGGGTTACTGATGATCCAAACATGTCTTGCAAGAG 1015  
 QY 157 ctggagaccacactctgttaacctctgttaccctgttccatgttaactgttaaccaagaca 216  
 DB 1016 CTGGGATCCATACGCTAGTGAATCTTGCACATGTTCATGTCATGCAACACAGAA 1075  
 QY 217 ccgcgtcactcgtgtga 234  
 DB 1076 CAGTGTCTATAAGAGTGTGA 1093

## RESULT 14

ID V14518 standard; DNA; 3979 BP.

AC V14518;

DT 20-MAY-1998 (first entry)

DE CF-5 pathogen resistance gene variant #1.

KM Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;

KM tomato leaf mould; Phytophthora resistance; ss.

OS Lycopersicon pimpinellifolium.

FR Key Location/Qualifiers

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

WO9743429-A1.  
 PD 20-NOV-1997.  
 PR 08-MAY-1997; G01249.  
 PR 24-SEP-1996; GB-019924.  
 PR 09-MAY-1996; GB-009681.  
 PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.  
 PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDS;  
 DR WPI: 98-008895/01.  
 DR P-PSDB; W41309.  
 PT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful  
 PT for production of transgenic plants resistant to pathogens e.g.  
 PT tomato leaf mould C. fulvum in tomatoes  
 PS Claim 3: Fig 1a; 75pp; English.  
 CC This sequence is an example of the polynucleotide of the invention, and  
 CC is able to confer pathogen resistance on a plant. It is one of two tomato  
 CC CF-5 gene variants, which offer resistance against the pathogen  
 CC Cladosporium fulvum-5. Transgenic plants can be produced by incorporating  
 CC the gene into plant cells and regenerating plants from the cells;  
 CC asexually or sexually produced offspring can also be subsequently



CC produced. Expression of the gene in plant cells can confer pathogen  
CC resistance on a plant e.g. to tomato leaf mould (*C. fulvum*) in tomatoes.  
CC oligonucleotides with sequences complementary to the gene or fragments  
CC of it, are useful in anti-sense techniques to reduce gene expression. The  
CC nucleic acids/polynucleotides are useful as hybridisation probes to  
CC identify other genes/fragments conferring pathogen resistance on plants  
CC e.g., Phytophthora resistance in potatoes. Homologies between Cf-5 and  
CC Cf-9 may be used to identify further resistance genes of this class.  
50 Sequence 3979 bp; 1217 A; 711 C; 658 G; 1347 T.

Query Match	8.9%	Score 70.6	DB 1	Length 3979
Best Local Similarity	50.4%	Pred. No. 1.6e-11		
Matches 172; Conservative	0	Mismatches 169	Indels 0	Gaps 0

Qy	215	aacgcgcactcgtgtagctttggggaattcaaacctctcggacatctggcgcttag	274
Db	13444	AACAACCTGTTTATGTGTTGATCTTTACAAATACAGCTTTCTGGCTCATTCCTTAAGA	20033
Qy	275	ctgggaagcttgaacatttaacgtacttagagctctacaanaacaacatccaaggaat	334
Db	2004	ATAGGTACCTGAGTTCTCTTACTGACATATTTTGGTAATAACCTCTTAAATGGCCT	20633
Qy	335	ataccttcgcgaacttggaaatctcgaagaatctcatagcttgatcttgcatacacaaat	394
Db	2064	ATTCTCTCTCATTTGGGGAACTCAACCACTTCTAGCTTGTATCTTTACAAATTAAG	21233
Qy	395	cttaagaagtagtcccaactcttttgggaaatbgaagctctcgtgcttttaagcgt	454
Db	2124	CTTTCCTGGCTATTCCTCGTTTCATTTGGCAAAATGAGAACTGCAAACTCTGTTC	21833
Qy	455	aatgaaccgatattgaaggggccaatccccagaagcactacatgcgaatcccaagccttaa	514
Db	2184	AGATATACGATCTCATTTGGGAAATTCCTCATTTGTGTGCAATTTGACATCGA	22433
Qy	515	gtgttgatgtctcaagaatgatgtgtgtgaaacatccc	555
Db	2244	GTTGTTATATGTGAGAAACAATTTTGAAGGAAAGTTCC	2284

**RESULT 15**

ID V14519 standard; DNA; 3979 BP.

DT	20-MAY-1998 (first entry)
DE	CF-5 pathogen resistance gene variant #2.
KW	Tomato; CF5 pathogen resistance gene; Cladosporium fulvum-5
OS	tomato leaf mould; Phytophthora resistance; ss.
MS	Lycopersicon pimpinellifolium.
FT	Location/Qualifiers
CDS	653..3560

PN WC09743429-A1.  
PD 20-NOV-1997.  
PE 08-MAY-1997: G01249.  
PR 24-SEP-1996: GB-019924.  
PR 09-MAY-1996: GB-009681.  
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.  
FI Dixon WS, Hatzixanthis K, Jones DA, Jones JDG;  
DR WPI: 98-008895/01.  
DR P-PSDB; W41310.  
PT Tomato gene Cf-5, confers resistance to *Cladosporium fulvum* - useful  
PT for production of transgenic plants resistant to pathogens e.g.  
PT tomato leaf mold *C. fulvum* in tomatoes  
PS Claim 6; Fig 1b; 75pp. English.  
CC This sequence is an example of the polynucleotide of the invention, and  
CC is able to confer pathogen resistance on a plant. It is one of two tomato  
CC Cf-5 gene variants, which offer resistance against the pathogen  
CC *Cladosporium fulvum*-5. Transgenic plants can be produced by incorporating  
CC the gene into plant cells and regenerating plants from the cells;  
CC asexually or sexually produced offspring can also be subsequently  
CC produced. Expression of the gene in plant cells can confer pathogen

CC resistance on a plant e.g. to tomato leaf mould(C. fulvum) in tomatoes.  
CC Oligonucleotides with sequences complementary to the gene or fragments  
CC of it, are useful in anti-sense techniques to reduce gene expression. The  
CC nucleic acids/polynucleotides are useful as hybridisation probes to  
CC identify other genes/fragments conferring pathogen resistance on plants  
CC e.g. Phytophthora resistance in potatoes. Homologues between Cf-5 and  
CC Cf-9 may be used to identify further resistance genes of this class.  
SD Sequence 3975 bp, 1218 A; 716 C; 698 G; 1347 T;

Query Match	8.9%;	Score 70.6;	DB 1;	Length 3979;
Best Local Similarity	50.4%;	Pred. No. 1.6e-11;		
Matches 172;	Conservative	0;	Mismatches 169;	Indels 0;
				Gaps 0;

Qy	215	aacgcgctccactcgtcygagctttggggaattccaaacccctctggacatcttggcccttag	274
Db	1944	AACAACTGTTATGTTATGTTATCTTTACATATATCAGCCTTTCGGCTCATTCCTGAGAA	2003
Qy	275	ctcyggaagcttbaacatttaacagfatctagagctcttacaanaacacatccagaact	334
Db	2004	ATAGGTTACCTGAGTCTCTTACTGACATATTTTGGGTAATACCTCTTAAATGGCTCT	20633
Qy	335	ataccttcgcgaacttggaaatctgaagaatctcatcagcttgatctgtacacacacat	394
Db	2064	ATTCCTGCTCATTTGGGGATCTAAACAACCTGTCTAGTGTGTATCTTTACATATATAG	21233
Qy	395	cttcaaggagatagttcccactctcttttgggaaattgaagctctctgctcttlltaacgctt	454
Db	2124	CTTTCCTGGCTCATTCCTCGTTCTATTGGCAATATGAGAACTCTCAAACTGTGTTTC	21833
Qy	455	aatcacacgataltgaagcggccaatccctagacatctacatgcacatcccaagccttaa	514
Db	2184	AGTATATACGATGCAATTTGGGGAATTCCTCATTTGTGTGCAATTGTGACATCACTGGA	22433
Qy	515	gtgtgtgatgtctcaagcaatgattgtgttggacaacatccc	555
Db	2244	GTGTTGTATATGTGAGAAACAATTTTGAAGGGAAGAACTTCC	2284

```
Search completed: June 23, 2000, 22:53:17
Job time: 40593 sec
```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Comphen Ltd.

OK nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:45:22 ; Search time 274.21 Seconds  
(without alignments)  
374.013 Million cell updates/sec

Title: US-09-180-798-26

Perfect score: 789

Sequence: 1 tcgacaccagcgctcgcgca.....agtcaaaaaaaaaaaaaa 789

Scoring table: IDENTITY\_NUC

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued\_Patents.NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/5C\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/5D\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/6\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/6CTUS\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63.6	8.1	3921	4	US-08-567-375-3
2	63.6	8.1	3921	4	US-08-567-680A-3
3	63.6	8.1	5992	3	US-08-475-891A-3
4	58	7.4	2075	1	US-08-238-163-3
5	55.8	7.1	1554	4	US-08-587-680A-24
6	55.4	7.0	1058	1	US-08-238-163-1
7	55.4	7.0	6256	3	US-08-475-891A-1
8	55.4	7.0	6256	4	US-08-567-375-1
9	55.4	7.0	6256	4	US-08-567-680A-1
10	53	6.7	5733	3	US-08-473-553A-1
11	50.8	6.4	831	4	US-08-567-375-15
12	48.4	6.1	2880	4	US-08-666-271-1
13	48.4	6.1	3905	4	US-08-666-271-1
14	41.4	5.2	1116	2	US-08-244-646-14
15	41.4	5.2	2917	2	US-08-592-936B-20
16	41.4	5.2	2917	4	US-09-111-573-20
17	40.6	5.1	792	2	US-08-244-646-16
18	40.6	5.1	792	2	US-08-442-063A-32
19	33	4.2	426	2	US-08-442-063A-35
20	33	4.2	567	2	US-08-442-063A-38
21	33	4.2	711	2	US-08-442-063A-41
22	33	4.2	849	2	US-08-442-063A-44
23	33	4.2	924	2	US-08-442-063A-47
24	33	4.2	1002	2	US-08-442-063A-26
25	33	4.2	1026	1	US-08-272-919-1
26	33	4.2	1026	2	US-08-619-916-1
27	33	4.2	1026	6	PCT-US95-08542-1

28	33	4.2	1593	7	US-08-480-346-1	Patent No. 5340934
29	31.2	4.0	2401	2	US-08-243-541-1	Sequence 1, Appl
30	31.2	4.0	2401	3	US-08-480-344-1	Sequence 1, Appl
31	31.2	4.0	2401	3	US-08-480-344-1	Sequence 1, Appl
32	30.6	3.9	3850	3	US-08-371-377-22	Sequence 22, Appl
33	30.2	3.8	2241	6	PCT-US95-10509-1	Sequence 1, Appl
34	30.2	3.8	4203	3	US-08-866-757-1	Sequence 1, Appl
35	30	3.8	1607	2	US-08-446-908-1	Sequence 1, Appl
36	30	3.8	1607	2	US-08-231-205A-1	Sequence 1, Appl
37	30	3.8	1607	4	US-08-871-161-1	Sequence 1, Appl
38	30	3.8	2088	4	US-08-602-264A-1	Sequence 1, Appl
39	30	3.8	6243	4	US-09-056-075-1	Sequence 1, Appl
40	29.8	3.8	457	3	US-08-189-256A-16	Sequence 16, Appl
41	29.8	3.8	1448	3	US-08-189-256A-11	Sequence 11, Appl
42	29.8	3.8	2363	1	US-08-096-947-2	Sequence 2, Appl
43	29.8	3.8	2363	2	US-07-919-140B-2	Sequence 2, Appl
44	29.8	3.8	2363	6	PCT-US93-06939-2	Sequence 2, Appl
45	29.6	3.8	594	1	US-08-289-458-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-567-375-3  
; Sequence 3, Application US/08567375  
; Patent No. 5952485  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Confering  
TITLE OF INVENTION: Disease Resistance in Plants  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/567,375  
FILING DATE: 04-DEC-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,645  
FILING DATE: 29-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/475,891  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,375  
FILING DATE: 17-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastien, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-058930  
TELEPHONE: (415) 576-0300  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear



```

? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: join(1..2676, 3520..3918)
? OTHER INFORMATION: /product="Xa-21"
US-08-567-375-3
```

Query Match	8.1%	Score 63.6;	DB 4;	Length 3921;
Best Local Similarity	50.3%;	Pred. No. 7.1e-11;		
Matches 156; Conservative	0;	Mismatches 154;	Indels 0;	Gaps 0;

[illegible]

RESULT 2  
US-08-587-680A-3  
: Sequence 3, Application US/08587680A  
: Patent No. 5977434  
: GENERAL INFORMATION:  
: APPLICANT: Ronald, Pamela C.  
: APPLICANT: Wang, Guo-Liang  
: APPLICANT: Song, Wen-Tuang  
: APPLICANT: Szabo, Veronique  
: TITLE OF INVENTION: Procedures and Materials for  
: Conferring Disease Resistance in Plants  
: NUMBER OF SEQUENCES: 27  
: CORRESPONDENCE ADDRESS:  
:

```

1      FILING DATE: 29-SEP-1995
2      PRIOR APPLICATION DATA:
3      APPLICATION NUMBER: US 08/567,375
4      FILING DATE: 04-DEC-1995
5      ATTORNEY/AGENT INFORMATION:
6      NAME: Bastian, Kevin L.
7      REGISTRATION NUMBER: 34,774
8      REFERENCE/DOCKET NUMBER: 023070-05894005
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: (415) 576-0200
11     TELEFAX: (415) 576-0300
12     INFORMATION FOR SEQ ID NO: 3:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 3921 base pairs
15     TYPE: nucleic acid
16     STRANDEDNESS: single
17     TOPOLOGY: linear
18     MOLECULE TYPE: DNA (genomic)
19     FEATURE:
20     NAME/KEY: CDS
21     LOCATION: join(1..2676, 3520..3918)
22     OTHER INFORMATION: /product= "xa-21"
23
24     US-08-567-680A-3

```

Query Match	8.18;	Score 63.6;	DB 4;	Length 3921;
Best Local Similarity	50.38;	Pred. No. 7.1e-11;		
Matches 156; Conservative	0;	Mismatches 154;	Indels 0;	Gaps 0;

Oy	248	aaatcctctgaaacctcttgagccttgagcttggaaagcttggaaacttgaacattcaagatctagag	307
Db	1159	AAAGTACAGGAAACCACTTCGAGAGGAAATTTGGCAATCTTATTTGGCTTAAACATCTCAT	1218
Oy	308	ctctacaaaacaacatccagaaggaaactataccttcaccaattggaatctgaaagatctc	367
Db	1219	CTTCGCAACAAACAATTTCACAGAGGGTCTCTCCATCTCTGGCAGGGTTAAAACTTA	1278
Oy	368	atcagcttgagctcgttacaacacaaactctacagagatagttcccaactcttctggaaaa	427
Db	1279	GGCATTTCTACGCGCTTACGAAAAACAATTGAGCGGTTTCAGTCCCGTTGGCCATAGGAAT	1338
Oy	428	ttgaagtccttgatccttcttaagcgttaatgaaacccgattgaagggccaatcccttga	487
Db	1339	CTTACTGAACTTAATATCTTACTTACCTGCGCACCAAAATTAAGTGTGGATTACCATAC	1398
Oy	488	gcaactcactgcaatcccaacagccttaagttgltgatagttccaaagcaatgatattgttga	547
Db	1399	ACACCTCAAAACCTCACAAACTGTGTGTCTATTAGGCGTTTCAACTTAATACCTTATGTGT	1458
Oy	548	acaatcccaaa	557
Db	1459	CCAAATCCCA	1468

RESULT 3  
US-08-475-891A-3  
Sequence 3, Application US/08475891A  
Patent No. 5859339  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-liang  
APPLICANT: Song, Wen-yuang  
TITLE OF INVENTION: Procedures and Materials for Confering  
TITLE OF INVENTION: Disease Resistance in Plants  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475.891A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373.375  
FILING DATE: 17-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 02370-0589100S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ. ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5992 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(512..3149, 3993..4393)  
OTHER INFORMATION: /product- "RRK-B"  
OTHER INFORMATION: /note- "Xaz1 Xanthomonas spp. disease  
OTHER INFORMATION: resistance gene RRK-B from rice (Oryza  
US-08-475-891A-3

Query Match 8.1%; Score 63.6; DB 3; Length 5992;  
Best Local Similarity 50.3%; Pred. No. 9.1e-11;  
Matches 156; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 248 aacctctctgacatcttgctgagcttgagcttgagcaattacagatctagag 307  
DB 1670 AAGATCAGAGGAGGATTCGAGGATATGCGCATCTTATGGCTTAAACATCTCAT 1729  
QY 308 ctctcaaaaacaacatccaggaactataccttccggaacttgaagaatctc 367  
DB 1730 CTCGCAACAACATTCAGAGGCTCTTCATCATCGTGGCGAGGCTTAAAACTTA 1789  
QY 368 atcagcttgatctgtacacaacaatcttacagggatagttccacttcttggaaa 427  
DB 1790 GGCATTTACTCGGCTACGAAACACTTGAAGGCTTCGATCCCTTGGCCATAGGAAT 1849  
QY 428 ttgaagctctgltcttttaaggcttaataagcaaccgattgacggggcaatccctaga 487  
DB 1850 CTTACTGAACTTATATCTTACTGCTCGGCAACAATAATTCAGTGTGATACCATAC 1909  
QY 488 gcaatcgaatcccaaccccttaagttgttgatgctcgaagcaatgattgtgtga 547  
DB 1910 ACACTTCGAACCTCAAAACTTGTGTCTATGAGGCTTCAACTAATTAACCTTAGTGT 1969  
QY 548 acaatcccaa 557  
DB 1970 CCAATACCA 1979

RESULT 4  
US-08-238-163-3  
Sequence 3, Application US/08238163  
Patent No. 5569830  
GENERAL INFORMATION:  
APPLICANT: BENNETT, Alan  
APPLICANT: LABAYTCH, John M.  
APPLICANT: POWELL, Ann  
APPLICANT: STOFZ, Henrik

TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend Kourile and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/238.163  
FILING DATE: 03-MAY-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 2307E-540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ. ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 421..1401  
US-08-238-163-3

Query Match 7.4%; Score 58; DB 1; Length 2075;  
Best Local Similarity 50.5%; Pred. No. 3.3e-09;  
Matches 196; Conservative 0; Mismatches 185; Indels 7; Gaps 2;

QY 171 ttgtatctctgttaccgtgttccatgtacactgtacacaaacacacgcgtcaactcgt 230  
DB 576 TTCTGTTACTGTTAGTCAATTAATGTGACCGGAAAC---CAACCGATTAATGCTC 631  
QY 231 tggatttgggaattcaaacctctctgacatcttgagccttgagcttggaggttgaac 290  
DB 632 TCACCGTCTTCCAGGCAATATCTCGGCAATTCGCGACCGCTGGAGACTTCCAT 691  
QY 291 atttaagatctagaagctctaca--aaacaatccaaagaaactatcacttcgaac 347  
DB 692 ATCTCAAAACATTTGGAATTTCACTGTTACTCAATCTCACCGAAATTCACCTCCAA 751  
QY 348 ttgaaatctgaagatctcatcagcttgatctgtataaacaacatcttacagagtag 407  
DB 752 TTGCGAGCTCACAAATCTCAAAATGTTAAGGCTTAGCTTCACTTAACCTTACAGAGTCCGA 811  
QY 408 ttccacttcttgggaaattgaagctctgttcttttaagcgttaatgacaacgat 467  
DB 812 TCCCTGAATCTTATGCTAGCTGAAGAAATTTGACGTTGCTGCGAGTTGATTAACAATCAAT 871  
QY 468 tgaaggggccaatccctaaagcaatcgaatcccaagccttaagttgttgatgtc 527  
DB 872 TTACCGGAACATCTCTTCTCCCTCTCACTTCCGAAATTTGCTAGCATGTACTTAC 931  
QY 528 caagcaatgattgttgtaaacatcc 555  
DB 932 ATGTAACAACTCACCGGAACAATAC 959

RESULT 5



```

US-08-587-680A-24
; Sequence 24, Application US/08587680A
; Patent No. 5977434
;
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yang
; APPLICANT: Szabo, Veronique
;
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 27
;
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; City: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA (partial)
;
US-08-587-680A-24

```

	Query Match	7.1%	Score 55.8	DB 4	Length 1554
	Best Local Similarity	49.5%	Pred.No. 1.5e-06		
	Matches 144	Conservative	0	Mismatches 147	Indels 0
				Gaps 0	
QY	250	cctctcgcgcacatctgcgcctgcgcttgagcttggagagcttgaacattacagatctcagaagct	309		
Db	18	CTTCTCTCGTGGACACTTCCTAGTGCCTATTGTGAAACATATTCAGGGCGAAGATCTTGCTT	77		
QY	310	ctacaaacacacatcccaagaaactatcactctccgaacttggaaattctgaagaattcat	369		
Db	78	AACGTGAAATGGTTTCCACAGTGATATCCCTCTGATATGTGCAGACATAAGACAACTT	137		
QY	370	cagcttggatctgtgacaacaacatcttcaagggatgtccactctcttgggaaatt	429		
Db	138	AAAGCTGACCTGTGTGAAACAACTTCTGTGGCAATCCCTCTCAGATGGTACTG	197		
QY	430	gaagctctctgctcttthtacgcgttaatgacaacagcatggaaggggccaatccctagaagc	489		

Db 198 TCTTTCTTACTACTACTTGATTTGAGCCAAATTCACATTTCTGGGCTATCCCAATTCA 257  
Qy 490 actcactgcatacccaagcccttaagttgttgaatgctcaagaatgatgtt 540  
Db 258 AATTGCTCAATTCACATCTTAAATTACATCATATTTTCTGGAAATCACTT 308

```

1      RESULT      6
2      US-08-238-163-1
3      ; Sequence 1, Application US/08238163
4      ; Patent No. 5569830
5      ; GENERAL INFORMATION:
6      ; APPLICANT: BENNETT, Alan
7      ; APPLICANT: LABAVITCH, John M.
8      ; APPLICANT: POWELL, Ann
9      ; APPLICANT: STORZ, Henrik
10     ; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
11     ; TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
12     ; NUMBER OF SEQUENCES: 24
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: Townsend and Townsend Hourie and Crew
15     ; STREET: Stewart Street Tower, One Market Plaza
16     ; CITY: San Francisco
17     ; STATE: California
18     ; COUNTRY: US
19     ; ZIP: 94105-1493
20     ; COMPUTER READABLE FORM:
21     ; MEDIUM TYPE: Floppy disk
22     ; COMPUTER: IBM PC compatible
23     ; OPERATING SYSTEM: PC-DOS/MS-DOS
24     ; SOFTWARE: Patent Release #1.0, Version #1.25
25     ; CURRENT APPLICATION DATA:
26     ; APPLICATION NUMBER: US/08/238,163
27     ; FILING DATE: 03-MAY-1994
28     ; CLASSIFICATION: 800
29     ; ATTORNEY/AGENT INFORMATION:
30     ; NAME: Bastian, Kevin L.
31     ; REGISTRATION NUMBER: 34,774
32     ; REFERENCE/DOCKET NUMBER: 2307E-540
33     ; TELECOMMUNICATION INFORMATION:
34     ; TELEPHONE: (415) 543-9600
35     ; TELEFAX: (415) 543-5043
36     ; INFORMATION FOR SEQ ID NO: 1:
37     ; SEQUENCE CHARACTERISTICS:
38     ; LENGTH: 1058 base pairs
39     ; TYPE: nucleic acid
40     ; STRANDEDNESS: single
41     ; TOPOLOGY: linear
42     ; MOLECULE TYPE: cDNA
43     ; FEATURE:
44     ; NAME/KEY: CDS
45     ; LOCATION: 34..1023
46     ; FEATURE:
47     ; NAME/KEY: misc feature
48     ; LOCATION: 1..1058
49     ; OTHER INFORMATION: /standard_name= "Pear PGIP cDNA"
50     US-08-238-163-1

```

Query Match	7.0%	Score 55.4	DB 1	Length 1058
Best Local Similarity	51.4%	Pred. No. 1.6e-08		
Matches 128	Conservative	0	Mismatches 121	Indels 0
			Gaps	0
QY	241	gaattcaaacctctctggagcatcttgcgctagcttgggaacttgaacatttaccagta	300	
Db	342	GCACCCCAATCTACTGCGCCCAATCCAAACCCGCCATTCGCAAGCTAAAGGACTCAAGTC	401	
QY	301	tctagaagcttacaacaaacacatccagaagaactatacttccgaacttggaaattgaa	360	
Db	402	TCTCAGGCTCAGCTGAGCACCACCTCTCAGAGCTCTGTCCCTGACTTCTCAGGCAACTCAA	461	
QY	361	gaattcattcattcagcttggatctgttacaacacaaactttacaggatattgtcccaattcttt	420	



Db 462 GAACTTACATTCCTCAGCTCTCTTCAACACCTCACCAGGCGCATCCAGCTCGCT 521  
Oy 421 gggaagaattgaagctctgctcttttaagcgttaatgcaaacgattgaagggaat 480  
Db 522 TTCTGAGCTCCCAAACTCGGCGCTTTCGTCTAGACCCGATTAAGCTCAGGTCAAT 581  
Oy 481 cccatgagc 489  
Db 582 TCCGATATC 590

RESULT 7  
US-08-475-891A-1  
; Sequence 1, Application US/08475891A  
; Patent No. 5859339  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; TITLE OF INVENTION: Procedures and Materials for Confering  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,891A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baetian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 02370-058910US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6256 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1648..4383, 5178..5513)  
; OTHER INFORMATION: /product="RRK-F"  
; OTHER INFORMATION: /note="Xa21 Xanthomonas spp. disease  
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza  
; OTHER INFORMATION: sativa)"  
US-08-475-891A-1

Query Match 7.08; Score 55.4; DB 3; Length 6256;  
Best Local Similarity 50.28; Pred. No. 4.5e-08;  
Matches 165; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

Oy 231 tggattgggaattcaaacctctctgacatcttgagccttgaggaagctgaac 290  
Db 2735 TGGACTTGGAGAAATTAATCACTGGGGGAGTCTCTTAATTCGTTTTCATCTTCCA 2794

Oy 291 att---taccgtatctagaagctctcaaaaacatccaaaggaactatacttcgaac 347  
Db 2795 CTTCGGTGAATTTCTCTGACCTTGATTTGAATGAATACAGGAAGCATTTCCAAAGATA 2854  
Oy 348 ttggaattctgaagaatctcaatcgaacttgatctgtatcaacaacaattctaaaggtag 407  
Db 2855 TTGGCAATCTTAATGGCTTACAAATCTCTATCTCTGCAACAAATTTCCAGAGGTCCAC 2914  
Oy 408 ttccactcttggaagaattgaagctctgctcttttaagcgttaatgacaacgat 467  
Db 2915 TTCCATCATCTGTTGGCGAGCTTGAAGAACTTAGGCATTTAGTCCACCAAAACACT 2974  
Oy 468 tgaaggggccaatccctagagcactcaactgcaatcccaagccttaagttgagtct 527  
Db 2975 TGAAGGTTGATGATCCATTCATAGGAATCTTACTGAACTTAATATTCTACTGCTCG 3034  
Oy 528 caagcaatgattgtgtggaacatccca 556  
Db 3035 GCACCAACAATTCAGTGGTTGATACCA 3063

RESULT 8  
US-08-567-375-1  
; Sequence 1, Application US/08567375  
; Patent No. 5952485  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Confering  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/567,375  
; FILING DATE: 04-DEC-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004,645  
; FILING DATE: 29-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475,891  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baetian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 02370-058930  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6256 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)







```

; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RET/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
;
; US-08-473-553A-1
;
Query Match
Best Local Similarity 6.7%; Score 53; DB 3; Length 5733;
Matches 128; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
;
QY 310 ctcaaaacaacatcccaaggaactatccgaacttcgaatcgaatccat 369
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3120 CTCTCTCAACACAGCTACACCGGTGGTGTTCACGAGCTTCGAGTTAAACAAAGCTTGA 3179
;
QY 370 cagcttgagatctgtacaacaacatctacaggaatgtccactctcttggaat 429
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3180 GATCCCTCGACATGGCGAGCTGTACACTCACCGAGAGATTCGAGAGATTAAAGTAACCT 3239
;
QY 430 gaagctctggtcttttcagcgcttaagaaacgcatgacggggccaatccctagaac 489
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3240 GAAACATCTACACTCTCTTTCTTCACATCAACAACTTAACCGGTCAATATCACCGGA 3299
;
QY 490 actcactgtaacccaagccttaagttgttgatgtctcaagcaatgattgttggaac 549
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3300 GCTTCCGGTTAGTCAGCTGAATTCCTCGATTATTCATCAATCAATGTAACCGGGA 3359
;
QY 550 aatcccaacaac 562
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3360 AATCCCTCAAGC 3372
;
;
RESULT 11
US-08-567-375-15
; Sequence 15, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants

```

```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; US-08-567-375-15
;
Query Match
Best Local Similarity 6.4%; Score 50.8; DB 4; Length 831;
Matches 139; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
;
QY 255 ctgagacatctggcgcctgagcttggaagcttgaacattacagatcttagagcttaca 314
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 33 CTGGTGCATTCCTTAGTCTTTTGAAACTATTCAGGGCTGAAGAACTGTGTAACTG 92
;
QY 315 aaaaacaacatcccaaggaactataccttcgaacttggaaatctgaagaatcactcaact 374
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 93 GAAAGGTTTCTCAGAGTATATCCCTCTGATATGCGACACTAAAGAGCATCTTAAGC 152
;
QY 375 tggactctgaacaacacatctacaggaatgagtcacactctcttggaatgaatgaat 434
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 TGGACTCGAGTAGAACAACCTTCTGTGCAAAATCCCTCTCAAGTTGGTAATGTCCTT 212
;
QY 435 ctctgtctcttttaacgcttaatgacaaccgatgacaggggccaatccctagaagcaacta 494
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 CCTTAACCTTACTTGATTTGAGCCAAATCAATCTTCTGCTCAATCCAGTTCAATG 272
;
QY 495 ctgcaatcccaagccttaagttgttgatgtctcaagcaatgatt 540
; ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 273 CTCAAATTCACATCTTAATTAATCAATATTTCTTGGAATCACTT 318
;
;
RESULT 12
US-08-666-271-4
; Sequence 4, Application US/08666271
; Patent No. 5920000
; GENERAL INFORMATION:

```







Db 2973 TGCACCTGACGTTGAACTTGTCTGCAATGCTGAAAGTCATATACCGCATCAT 3032  
QY 421 gggaaattgaagtcctgtcttttaagcgctaatgcaacccgattgacggggccaat 480  
Db 3033 TCANAAATTTATCGATACCTGCAATCTTGGATCTTCATCTAATAAATCAGCGGAGAAAT 3092  
QY 481 ccctagagcactcactgcaatcccaagccttaagttgtatgtctcacaagcaatgatt 540  
Db 3093 TCCGACAGCAGCTTGCAATCCCTCACAATTCCTTGAAAGTCTAATAATCTCTCACAATCATCT 3152  
QY 541 gtgtggacaatcccaacaacggaccctttgct 574  
Db 3153 TGTGTGATGCATCCCAAGAAAACAAATTTGAT 3186

## RESULT 14

US-08-244-646-14  
Sequence 14, Application US/08244646  
Patent No. 5744692

## GENERAL INFORMATION:

APPLICANT: Cervone, Felice  
APPLICANT: De Lorenzo, Giulia  
APPLICANT: Salvi, Giovanni  
APPLICANT: Alberghini, Peter  
APPLICANT: Darvill, Alan  
TITLE OF INVENTION: Nucleotide Sequences Coding An  
NUMBER OF INVENTION: Endopolygalacturonase Inhibitor  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sally A. Sullivan  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,646  
FILING DATE: 06-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT RM 91A 000915  
FILING DATE: 06-DEC-1991  
PRIOR APPLICATION DATA: WO PCT/IT/00158  
APPLICATION NUMBER: WO PCT/IT/00158  
FILING DATE: 04-DEC-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Sullivan, Sally A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 19-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1116 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Phaseolus vulgaris  
STRAIN: Saxa  
IMMEDIATE SOURCE:  
CLONE: lambda pGIP-3.3

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1026  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 1027..1116  
US-08-244-646-14

Query Match 5.2%; Score 41.4; DB 2; Length 1116;  
Best Local Similarity 48.1%; Pred. No. 0.00062;  
Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 313 caaaacacatcccaaggaactatacctccgaacttggaaatcgaatcgcacg 372  
Db 342 CATCAATTAACCTGTCGTCGCAATCCCGCCATCTGTAATCTACCCACCACTCA 401  
QY 373 ctggatctgtaacaacatctacagggatgctccactctcttgggaatgaa 432  
Db 402 TCTCTATACCTACACACCAATGTCCTCGCGCAATACCCATTTCTGTACAGATCA 461  
QY 433 gtctcgtctcttttaagcgttaataagacgattgacggggccaaaccctagagact 492  
Db 462 AACCTCGTACCCCTGCACTTCTCTACACGACCCCTCTCGGCAACCCCTCCCTCAT 521  
QY 493 cactgaatcccaagccttaagttgtatgtctcacaagcaatgattgtgtgaaat 552  
Db 522 CTCTCTCTCCCAACCTCGAGAAATCATTCGACGCAACCGAATCTCGGCGCCAT 581  
QY 553 CCC 555  
Db 582 CCC 584

## RESULT 15

US-08-592-936B-20  
Sequence 20, Application US/08592936B  
Patent No. 5783393

## GENERAL INFORMATION:

APPLICANT: Kelloeg, Jill A.  
APPLICANT: Bestwick, Richard K.  
TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR  
NUMBER OF INVENTION: REGULATED EXPRESSION OF TRANSGENES IN PLANTS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,936B  
FILING DATE: 29-JAN-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Evans, Susan T.  
REGISTRATION NUMBER: 38,443  
REFERENCE/DOCKET NUMBER: 4257-0012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 324-0880  
TELEFAX: (650) 324-0960  
INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2917 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:35 ; Search time 6198.48 Seconds

(without alignments)  
515.932 Million cell updates/sec

Title: US-09-180-798-26

Perfect score: 789  
Sequence: 1 tcgaccacgcgcgcgcgaa.....agtccaaaaaaaaaaaaa 789

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: em\_est10:\*  
11: em\_est11:\*  
12: em\_est12:\*  
13: em\_est13:\*  
14: em\_est14:\*  
15: em\_est15:\*  
16: em\_est16:\*  
17: em\_est17:\*  
18: em\_est18:\*  
19: em\_est19:\*  
20: gb\_est1:\*  
21: gb\_est2:\*  
22: gb\_est3:\*  
23: gb\_est4:\*  
24: gb\_est5:\*  
25: gb\_est6:\*  
26: gb\_est7:\*  
27: gb\_est8:\*  
28: gb\_est9:\*  
29: gb\_est10:\*  
30: gb\_est11:\*  
31: gb\_est12:\*  
32: gb\_est13:\*  
33: gb\_est14:\*  
34: gb\_est15:\*  
35: gb\_est16:\*  
36: gb\_est17:\*  
37: gb\_est18:\*  
38: gb\_est19:\*  
39: gb\_est20:\*  
40: gb\_est21:\*  
41: gb\_est22:\*  
42: gb\_est23:\*  
43: gb\_est24:\*  
44: gb\_est25:\*

45: gb\_est26:\*  
46: gb\_est27:\*  
47: gb\_est28:\*  
48: gb\_est29:\*  
49: gb\_est30:\*  
50: gb\_est31:\*  
51: gb\_est32:\*  
52: em\_est20:\*  
53: em\_est21:\*  
54: em\_est22:\*  
55: em\_est23:\*  
56: em\_est24:\*  
57: em\_est25:\*  
58: em\_est26:\*  
59: gb\_est33:\*  
60: gb\_est34:\*  
61: gb\_est35:\*  
62: gb\_est36:\*  
63: gb\_est37:\*  
64: gb\_est38:\*  
65: em\_est27:\*  
66: em\_est28:\*  
67: em\_est29:\*  
68: em\_est30:\*  
69: gb\_est39:\*  
70: gb\_est40:\*  
71: gb\_est41:\*  
72: gb\_est42:\*  
73: gb\_est43:\*  
74: gb\_est44:\*  
75: em\_est31:\*  
76: em\_est32:\*  
77: em\_est33:\*  
78: em\_est34:\*  
79: gb\_est45:\*  
80: gb\_est46:\*  
81: gb\_est47:\*  
82: gb\_est1:\*  
83: gb\_est2:\*  
84: gb\_est3:\*  
85: gb\_est4:\*  
86: em\_est1:\*  
87: em\_est2:\*  
88: em\_est3:\*  
89: em\_est4:\*  
90: gb\_est5:\*  
91: gb\_est6:\*  
92: gb\_est7:\*  
93: gb\_est8:\*  
94: gb\_est9:\*  
95: em\_est5:\*  
96: em\_est6:\*  
97: em\_est7:\*  
98: em\_est8:\*  
99: em\_est9:\*  
100: em\_est10:\*  
101: em\_est11:\*  
102: gb\_est10:\*  
103: gb\_est11:\*  
104: em\_est12:\*  
105: gb\_est12:\*  
106: gb\_est13:\*  
107: gb\_est14:\*  
108: gb\_est15:\*  
109: gb\_est16:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES



No.	Score	Match	Length	DB	ID	Description
1	395.4	50.1	556	23	R89998	R89998 16353 Lambd
2	368.6	46.7	556	23	A1992868	A1992868 701493826
3	349.4	44.3	364	42	A1100683	A1100683 33804 Lam
4	332	42.1	348	42	A1100682	A1100682 33803 Lam
5	327.8	41.5	353	23	H37300	H37300 15429 Lambd
6	318.2	40.3	534	33	AA394359	AA394359 25942 Lam
7	317.2	40.2	608	81	AA443205	AA443205 EST308135
8	314.6	39.9	720	64	AM031110	AM031110 EST274417
9	314.6	39.9	339	42	AT100679	AT100679 33800 Lam
10	305.8	38.8	401	37	AA712221	AA712221 31949 Lam
11	301.4	38.2	336	23	H37296	H37296 15425 Lambd
12	301	38.1	466	74	AM221939	AM221939 EST296750
13	297.6	37.7	657	80	AM50720	AM50720 GM210009A
14	296.8	37.6	511	74	AM221278	AM221278 EST297747
15	293.4	37.2	714	64	AM030188	AM030188 EST273443
16	281.6	35.7	334	23	H37195	H37195 15324 Lambd
17	280	35.5	685	64	AM038168	AM038168 EST279825
18	278.6	35.3	305	25	ME5416	ME5416 20456 Lambd
19	276	35.0	599	63	AM011134	AM011134 ST17B03 P
20	274	34.7	599	74	AM219797	AM219797 EST302279
21	265.4	33.6	637	80	AM50549	AM50549 GM210009A
22	264.4	33.5	485	45	AI352795	AI352795 MB61-10D
23	261.4	33.1	676	74	AM220075	AM220075 EST302558
24	257.6	32.6	430	79	AM307218	AM307218 sf54c07 Y
25	257.4	32.6	559	59	AT775448	AT775448 EST256548
26	253.8	32.2	447	23	H36800	H36800 14929 Lambd
27	246	31.2	545	64	AM037836	AM037836 EST279465
28	242	30.7	522	46	AI441759	AI441759 sa82d08 Y
29	237.4	30.1	667	74	AM185847	AM185847 se60d04 Y
30	236.2	29.9	336	20	Z34187	Z34187 ATTS3221 Ve
31	235.2	29.8	541	64	AM036865	AM036865 614019G10
32	233.6	29.6	619	64	AM040882	AM040882 EST283442
33	232.8	29.5	612	47	AI96525	AI96525 sb05c09 Y
34	228.4	28.9	423	35	C22371	C22371 C22371 R1ce
35	225.8	28.6	570	50	AI676939	AI676939 605047A07
36	225.2	28.5	515	79	AM79515	AM79515 sf90e09 Y
37	222.6	28.2	568	47	AI487272	AI487272 EST245594
38	221.4	28.1	616	42	AF074734	AF074734 AF074734
39	219.8	27.9	540	64	AM033599	AM033599 EST277170
40	218.2	27.7	497	59	AT772079	AT772079 EST253179
41	213.2	27.0	488	64	AM031917	AM031917 EST275671
42	210.6	26.7	463	74	AM221200	AM221200 EST297669
43	209.8	26.6	442	42	AI100481	AI100481 34856 Lam
44	200.2	25.4	337	20	Z34606	Z34606 ATTS3366 Ve
45	196.8	24.9	360	42	A1100685	A1100685 33806 Lam

## ALIGNMENTS

RESULT 1  
LOCUS R89998 556 bp mRNA  
DEFINITION 16353 Lambda-PR2 Arabidopsis thaliana cDNA clone 187H577, mRNA  
sequence.  
ACCESSION R89998  
VERSION R89998.1 GI:957538  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsis.  
REFERENCE 1 (bases 1 to 556)  
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.  
TITLE Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
JOURNAL Plant Physiol. 106, 1241-1255 (1994)

MEDLINE 95148729  
COMMENT On Apr 14, 1993 this sequence version replaced gi:693023.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PR2, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313c@ldm.cl.msu.edu  
Seq primer: r7 dye primer.  
location/Qualifiers  
1..556  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="187H577"  
/note="Vector: lambda Zip-lox; Site.1: Sal; Site.2: Not;  
lambda PR2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
silicles). The vector is BRU's lambda Zip-lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dr primed cDNA."

## FEATURES

BASE COUNT 151 a 118 c 123 g 146 t 18 others  
ORIGIN  
Query Match 50.1%; Score 395.4; DB 23; Length 556;  
Best Local Similarity 94.4%; Pred. No. 1,1e-98;  
Matches 441; Conservative 0; Mismatches 20; Indels 6; Gaps 3;  
QY 251 ctctctgacatctctgcccctgagcttggaaagcttgaacattcaagtaagctc 310  
DB 1 CTCCTGACATCTTGGCCCTGAGCTTGGAAAGCTTGAACATTATACGATCTAGAC 60  
QY 311 tacaataaacatccaaaggaactatcccttcgaacttggaaatcgaagatccac 370  
DB 61 TACAAACCAACATCCAAAGAACTATACCTTCCGAACCTGGAATCGAAGATCTCATC 120  
QY 371 agcttgagctctgtaacaacaacattacagagatgtccacatctcttgggaaatg 430  
DB 121 AGCTTGAGCTCTGTAACAACAATCTTACAGAGATGATTCACCTGTTGGGAAATTG 180  
QY 431 aagtcctgtgtcttcttaaggcttaatgaacaacagattgacggggccaatccctaga 480  
DB 181 AAGTCTGTGTCTTCTTACGGGCTTAATGACMACCGATTGACGGGCGCAATCCCTAG 240  
QY 491 ctcaactgaatcccaagccttaagttgttgatgtctccaaagcaatgtgtgtgaca 550  
DB 241 CTCACCTGCAATCCCAAGGCTTAAAGTTGTATGATCTCAAGCAATGATGTGTGACA 300  
QY 551 atcccaacaacagaccttctgtcacatcttccttaagaacttggaaacaaacaggag 610  
DB 301 ATCCCAACAACAGGACCTTTGTGCATTCCTTACAGAACTTGGANCAACCCGAGG 360  
QY 611 ttgg-agggacccgaatctacatcggtcttgcagactgaagacactacgcaccgaa 669  
DB 361 TTGGNAGGACCGGAAATTAATCGGCTTGAAGGTTCGCGCAATGCGCAATGCGAAC 417  
QY 670 ttggcaaaacctgaatgaagaaattggggggtgactcttgaagaac 716  
DB 418 ATTGGAACCTGAAA--TGAGGATGGGGGCTGNCCTTTAAGGCAC 462

RESULT 2  
LOCUS A1992868 469 bp mRNA  
DEFINITION 701493826 A. thaliana, Ohio State clone set Arabidopsis thaliana



ACCESSION CDNA clone 701493826, mRNA sequence.  
 A1992868  
 VERSION A1992868.1 GI:5839773  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsids.  
 1 (bases 1 to 469)  
 REFERENCE  
 AUTHORS Chen, J., Montiyama, M., Chan, E., Mooney, M., Carron, B., Gilliland, D.,  
 Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,  
 Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,  
 Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C.,  
 Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,  
 Norrige, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and  
 Hanson, D.  
 Arabidopsis thaliana Gene Expression MicroArray  
 Unpublished (1999)  
 On Dec 20, 1995 this sequence version replaced gi:1135328.  
 COMMENT  
 CONTACT: David Smoller, Ph.D.  
 Genome Systems, Inc., a wholly owned subsidiary of Incyte  
 Pharmaceuticals, Inc.  
 4633 World Parkway Circle, St. Louis, MO 63134, USA  
 Tel: 877-577-2733  
 Fax: 314-427-3324  
 Email: service@genomesystems.com.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..469  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /clone="701493826"  
 /note="CDNA library was made from selected clones from the  
 Arabidopsis thaliana Ohio State clone set."  
 BASE COUNT 120 a 124 c 89 g 136 t  
 ORIGIN  
 Query Match 46.7%; Score 368.6; DB 63; Length 469;  
 Best Local Similarity 97.7%; Pred. NO. 2.7e-91;  
 Matches 3/4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 5 cccacgcgcgcgcgaactatcgctgagagctcttcgcagcttcgttaacttaactta 64  
 87 CACATGGCTCTCGAAACTATCGGTGGAGCTCTCGCAGCTGTGTTATCTTAACCTTA 146  
 QY 65 gctttatcaccctggtcgagcaaacctccgagagagatgctcttaacgctctcgccgg 124  
 147 GCTTTATTCACCTGGTCCAAAGCAATCCGAAAGAGATGCTTTTACGCTCTTCGCCGG 206  
 QY 125 agttaacagatccgagacatgtctccagagctgagatccaactctgttaactcctgt 184  
 207 AGTTTAAAGATCCGAGCATGTCTCCAGACTGGGATCCAACTCTGTAAATCCTTGT 266  
 QY 185 acctggttcacgtcaactgtatacaagaacacgcgcgtcaactcgtgagatcttgggaat 244  
 267 ACCTGTTCCATGTACCTGTAAACCAAGCAACCCGCTACTCGTGTGATTTGGGGAAT 326  
 QY 245 tcaaacctctcggacatcttgccctgagcttggaagcttgaaacttaacttaactta 304  
 327 TCAAACTCTCTGAGCATTTGGCTGAGCTTGGGAAGCTTGAACATTTACATATCTTA 386  
 QY 305 gagctcacaacaaacacacccaaggaactatacctccgaacttggaactcgaagaat 364  
 387 GAGCTCTTAAATTAACATCCAAAGAACTATACCTCCGAACTTGAAATCTGAAGAT 446  
 QY 365 ctcatcagcttgatctgtaca 387  
 447 CTCATCAGCTTGATCTGTACAA 469  
 Db

RESULT 3  
 A1100683  
 LOCUS A1100683  
 DEFINITION 33804 Lambda-PRL2 Arabidopsis thaliana CDNA clone 187H5XP 3', mRNA  
 sequence.  
 ACCESSION A1100683  
 VERSION A1100683.1 GI:3449393  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsids.  
 1 (bases 1 to 364)  
 REFERENCE  
 AUTHORS Newman, T., deBrujin, F. J., Green, P., Keegstra, K., Kende, H.,  
 McIntosh, L., Ohlrogge, J., Ralkehl, N., Somerville, S., Thomasnow, M.,  
 Retzel, E. and Somerville, C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148725  
 On May 5, 1995 this sequence version replaced gi:797711.  
 COMMENT  
 CONTACT: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313tcn@lim.ci.msu.edu  
 The sequence entry for this EST has been reverse complemented and  
 is being submitted in the sense orientation.  
 Seq primer: M13-21.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..364  
 /organism="Arabidopsis thaliana"  
 /strain="var columbiana"  
 /db\_xref="taxon:3702"  
 /clone="187H5XP"  
 /note="Vector: Lambda Z1p-Lox; Site.1: Sal; Site.2: Not;  
 Lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRU's Lambda Z1p-Lox. The CDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dT primed cDNA."  
 BASE COUNT 113 a 80 c 70 g 99 t 2 others  
 ORIGIN  
 Query Match 44.3%; Score 349.4; DB 42; Length 364;  
 Best Local Similarity 98.9%; Pred. NO. 5.1e-86;  
 Matches 361; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 410 cccactctcttgggaactgaagctcgtctcttcttaagcgttaatgaacacgatg 469  
 1 CCCATTTCTTTGGGAAATTAAGTCCTNT-CTCTTTTACGGCTTAATGACACGATTTG 59  
 QY 470 acggggccaatccctagagcactcagcgaatcccaagccttaagtgttgatgtctca 529  
 60 ACGGGGCCAATCCNTAGAGCACTCAGTCATCCCAAGCCCTTAAGTGTGATGTCTCA 119  
 QY 530 agcaatgatgtgtggaacatcccaacaaaggaactttctcaacttcttaag 589  
 120 AGCAATGATTTGTGTGGAACAATCCCAAGGAGACTTTGTCTCATCTCTTTACAG 179  
 QY 590 aacttgagacacccgaggttgaggagccggaatctactcgtcttgcaagctagac 649  
 Db



Db 180 AACTTGAGAACACCGAGGTTGGAGGACCGGAATTACTCGCTTTCGCAACCTACGAC 239  
OY 650 actactgaccctgaaaaaatggcaaacctgaaatgagaattgggggtgacctg 709  
|||||  
Db 240 ACTAATCTACCTGCTAATAAATATGGCAAACTGAAATGAGAATTTGGGGGGGACCTTG 299  
OY 710 taagaacactcaccacttatcaatacacatctactatgttaataatgatatatgt 769  
Db 300 TAAGAACACTTACCACTTATCAAAATATCATCTACTATGTAAATAGTATATATGT 359  
OY 770 agtcc 774  
|||||  
Db 360 AGTCC 364

RESULT 4  
A1100682 348 bp mRNA EST 21-AUG-1998  
LOCUS 33803 lambda-PRL2 Arabidopsis thaliana cDNA clone 185LXP 3', mRNA  
DEFINITION  
ACCESSION A1100682  
VERSION A1100682.1 GI:3449392  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Arabidopsis thaliana  
thale cress.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsids.  
1 (bases 1 to 348)  
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)  
95148729  
On May 5, 1995 this sequence version replaced gi:797667.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcm@dm.cl.msu.edu  
The sequence entry for this EST has been reverse complemented and  
is being submitted in the sense orientation.  
Seq primer: M13-21.

FEATURES  
source  
1..348  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/strain="var columbiana"  
/db\_xref="taxon:3702"  
/clone="185LXP"  
/clone\_11b="lambda-PRL2"  
/note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not;  
lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques. The vector is BRU's lambda zip-lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dt primed cDNA. "

BASE COUNT 110 a 74 c 66 g 93 t 5 others  
ORIGIN

Query Match 42.1%; Score 332; DB 42; Length 348;  
Best Local Similarity 99.1%; Pred. No. 3, 2e-81;  
Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 440 gtcttttaaggcttaatgacacacgatgaggggccatccctagacactcaatgca 499  
|||||  
Db 14 GCTTTTACGCGCTTAATGACACCGAATTGACGGGGCCAACTCCATAGACATCAGCA 73  
OY 500 atcccaagccttaagtggtgtgatgctcgaagcaatgattgtgtggaacaatccaca 559  
|||||  
Db 74 ATCCCAAGCCTTAAGTGTGATGCTCAAGCAATGATTGTGTGCAAAATCCCAACA 133  
OY 560 aacggaccttgcctcacatcccttaacgaaccttgaagaacacccgaggttgaggga 619  
Db 134 AACGACCTTTTGCTCACATTCCTTACAGAACTTTGAGAAACCCGAGGTGGAGGA 193  
OY 620 ccggaattactcgtcttcgaagctagacactacgcacctgaataaattggcaaac 679  
Db 194 CCGGAATTAATCTGNTCTGCAAGCTAGCAGCACTACCTGCAAAAATTTGGCAAAAC 253  
OY 680 ctgaatatgaagaattgggggtgacctgtgaagaacactcaccacttatacaatc 739  
Db 254 CTGAATAATGAATAATTGGGGGTGANTCTGTAGAACACTGCACCACTTATCAAAATATC 313  
OY 740 acactactatgtaataagtatatatatgtagtc 774  
Db 314 ACATCTACTATGTATAATAGTATATATATAGTATGTC 348

RESULT 5  
H37300 353 bp mRNA EST 30-DEC-1997  
LOCUS H37300  
DEFINITION 15429 lambda-PRL2 Arabidopsis thaliana cDNA clone 179J197T, mRNA  
sequence.  
ACCESSION H37300  
VERSION H37300.1 GI:906799  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Arabidopsis thaliana  
thale cress.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsids.  
1 (bases 1 to 353)  
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)  
95148729  
On May 8, 1995 this sequence version replaced gi:801146.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcm@dm.cl.msu.edu  
Seq primer: T7 dye primer.

FEATURES  
source  
1..353  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/strain="var columbiana"  
/db\_xref="taxon:3702"  
/clone="179J197"  
/clone\_11b="lambda-PRL2"  
/note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not;  
lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques. The vector is BRU's lambda zip-lox. The cDNA



Inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA. " 14 others

Query Match 41.5%; Score 327.8; DB 23; Length 353;  
Best Local Similarity 94.1%; Pred. No. 4.6e-80;  
Matches 332; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

BASE COUNT 98 a 85 c 67 g 89 t  
ORIGIN  
Query Match 41.5%; Score 327.8; DB 23; Length 353;  
Best Local Similarity 94.1%; Pred. No. 4.6e-80;  
Matches 332; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
Query 311 tacacaaacacatccacgaagactatatactctcgaacttggaatctgaagaatctcacc 370  
Db 1 TACAAAAACAAATCCACGAGACATATCTCGAATTGGAATCTGAAGANTCTCATC 60  
Query 371 agcttgatctgtacacacacatctcaaggatagttccactctcttggaataattg 430  
Db 61 AGCTTGCTGTCTGACAAACAACTTACAGGATAGTCCACTTTTGGGAAATGG 120  
Query 431 aagctctgtctctttacagccttaagtacacacgattgacggggccacccctagaaga 490  
Db 121 AAGTCTCTGCTTTTACCGCTTAATGACACGATGACCGGTCCATCCCTAGAGCA 180  
Query 491 ctcaactgcaatcccaagccttaagtgttgatgtctcaagcaatgattgtgtgaaga 550  
Db 181 CTCACGGCAATCCCAAGCCTTAAGTGTGANTCTCANGCAATGATTTGTGTGAGACA 240  
Query 551 atcccaacaaacggacctttgtctcacatcccttaacagactttgaagaacccagg 610  
Db 241 ATCCCAACAAAGCNCCTTTGCTCACAATTCCTTACAGAACTTTGNGAACAACCGGNA 300  
Query 611 ttggaaggacgggaattactcgtctgtcgaagtcagacactactgacactg 663  
Db 301 TTGGAGGGGCGGANTTACTCGTCTGCAAGNTACGGCACTAATGCAACACTG 353

RESULT 6  
AA394359 534 bp mRNA EST 30-OCT-1997  
LOCUS 25942 Lambda-PRU2 Arabidopsis thaliana cDNA clone 30561T7 3', mRNA  
DEFINITION  
AA394359  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

On May 18, 1995 this sequence version replaced gi:811121.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313cne@lm.c1.msu.edu  
Seq primer: T7.

FEATURES  
source  
1..534  
/organism="Arabidopsis thaliana"  
/strain="var Columbia"  
/db\_xref="taxon:3702"  
/clone="30561T7"

/clone.lib="Lambda-PRU2"  
/note="vector: lambda zip-lox; Site.1: Sal; Site.2: Not;  
Lambda PRU2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark-rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
siliques). The vector is BRU's Lambda zip-lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA. "

Query Match 41.4%; Score 327; DB 33; Length 534;  
Best Local Similarity 88.0%; Pred. No. 8.1e-80;  
Matches 374; Conservative 0; Mismatches 45; Indels 6; Gaps 2;

BASE COUNT 151 a 114 c 125 g 126 t 18 others  
ORIGIN  
Query Match 41.4%; Score 327; DB 33; Length 534;  
Best Local Similarity 88.0%; Pred. No. 8.1e-80;  
Matches 374; Conservative 0; Mismatches 45; Indels 6; Gaps 2;  
Query 369 tcagcttgatctgtacacacacatctcaaggatagttccactctcttggaataat 428  
Db 1 TCAGCTTGATCTGTACACACATCTTACAGGATAGTCCACTTCTTGGGAAAT 60  
Query 429 tgaagctctgtctcttttcaagccttaagtcacacacgattgacggggccacccctagaag 488  
Db 61 TGAAGCTCTGTCTTTTACGGCTTAATGACACACGATGACCGGTCCATCCCTAGAG 120  
Query 489 caccactgcaatcccaagccttaagtgatgtctcaagcaatgattgtgtgaaga 548  
Db 121 CACTCAGGCAATCCCAAGCCTTAAGTGTGACCTTCAAGCAATGATTTGTGTGAAA 180  
Query 549 caatcccaacaaacggacctttgtctcacatcccttaacagactttgaagaacccaga 608  
Db 181 CAATCCCAACAAAGGACCTTTGCTCACAATTCCTTACAGAACTTTGAGAACCCGGA 240  
Query 609 ggttgaggagacgggaattactcgtctgtcgaagtcagacactactgacactgaa 668  
Db 241 GATTGGAGGAGACCGGAATTCCTGCTGCAAGCTACGACTAATGNACTGAACAA 300  
Query 669 attgcaaaccttggaatgaagattgggggtg---accctgaaagaacactcacca 725  
Db 301 ACTGCAAAACCTGGAATGAGAGNATTGGGGGGGTGACCTTTGTAAGACACTTCANCA 360  
Query 726 cttatacaatatactctac---tatgtaataagatatataatggtcacaataaa 782  
Db 361 CTTTATCCATATCCACATCCATTANGAATAGTATATAGTGTAAACCA 420  
Query 783 aaaaa 787  
Db 421 AAAAA 425

RESULT 7  
AM443205 608 bp mRNA EST 17-FEB-2000  
LOCUS EST308135 tomato mixed elicitor, B71 Lycopersicon esculentum cDNA  
DEFINITION  
AM443205  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

On May 18, 1995 this sequence version replaced gi:811121.  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313cne@lm.c1.msu.edu  
Seq primer: T7.  
Location/Qualifiers  
1..534  
/organism="Arabidopsis thaliana"  
/strain="var Columbia"  
/db\_xref="taxon:3702"  
/clone="30561T7"







Db 483 |||||CTGTACACACATATTTTCGGGACAAATCTCTACTCTTGGAACCTGAAAAACCTT 542  
 Qy 440 gtcttttaagcgttaatgacacacgattgacggggccatcccttagagcactcgtca 499  
 Db 543 GTTTCTTGCGTCTAATATGATACAGCTACAGACCAATCCCAAGAACTTACTGAC 602  
 Qy 500 atcccaagccttaagttgttgatgtctcaagcaatgatgtgtgtggaacaatccca 559  
 Db 603 ATTCTAGCCTGGAAGTGTGGATGTCTGGAATATGATTGTGTGGAACAATCTCTACT 662  
 Qy 560 aacggaccttgcctacattcctttacagaaacttgagaacacccgagttgagg 617  
 Db 663 TCTGGTCATTTCAGCATATTCCTCTAAACAAATTCGACACAAATCCTCGACTTGAAG 720

RESULT 9  
 A1100679 339 bp mRNA EST 21-AUG-1998  
 LOCUS 33800 Lambda-PRL2 Arabidopsis thaliana cDNA clone 179J11XP 3', mRNA  
 DEFINITION  
 ACCESSION A1100679 GI:3449389  
 VERSION A1100679.1  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.  
 1 (bases 1 to 339)  
 Newman,T., deBrujin,F.J., Green,P., Keegstra,K., Kende,H.,  
 McIntosh,L., Ohlrogge,S., Raikhel,N., Somerville,S., Thomashow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 On May 5, 1995 this sequence version replaced g1:797664.  
 CONTACT: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313tcn@dm.cl.msu.edu  
 The sequence entry for this EST has been reverse complemented and  
 is being submitted in the sense orientation.  
 Seq primer: M13-21.

## FEATURES

## source

1..339  
 /organism="Arabidopsis thaliana"  
 /strain="var columbiana"  
 /db\_xref="taxon:3702"  
 /clone="179J11XP"  
 /note="Vector: lambda zip-lox; Site\_1: Sal; Site\_2: Not;  
 Lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRV's lambda zip-lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dT primed cDNA." 89 t 5 others  
 BASE COUNT 104 a 75 c 66 g  
 ORIGIN

Query Match 39.9%; Score 314.6; DB 42; Length 339;  
 Best Local Similarity 97.0%; Pred. No. 2e-76;

Matches 328; Conservative 0; Mismatches 9; Indels 1; Gaps 1;  
 Qy 438 tggcttttaagcgttaatgacacacgattgacggggccaa-tcccttagagcactcact 496  
 Db 2 TGTCTTTTACGGCTTATGACAAACGATTCAGGGGGCAATCCCTAAGACACACAC 61  
 Qy 497 gcaatcccaagccttaagttgttgatgtctcaagcaatgatgtgtgtggaacaatccca 556  
 Db 62 GCATTGCCAAGCCTTAAAGTGTGTATNTCTCAAGCAATGATTTGTGTGGAACAAATCCCA 121  
 Qy 557 acaaacggaccttgcctacattcctttacagaaacttgagaacacccgagttgag 616  
 Db 122 ACAACGAGACCTTTGTCACATTCCTTTACAGAACTTGGAGAACACCCGAGTTGGAG 181  
 Qy 617 ggacgggaattcctcggtcttgcaagctagacactaactgacctgaaaaaattggcaa 676  
 Db 182 GGACCGGAATTACTCGGGCTTCGACATACACACTAATCGACCTGAAAAAATTTGGCAA 241  
 Qy 677 aacctgaatgaagaatgggggtgagccttgtaagaacacctcagcactttatcaat 736  
 Db 242 AAMCTGAATTAAGCAATTTGGGGGTGACCTTGTAAGAACCTTCACCACTTTAGCAAT 301  
 Qy 737 atcacatcctactgtaabaagtatataatagtgtcc 774  
 Db 302 AACCACATCTACTATGTAATATATATATATATAGTACNC 339

RESULT 10  
 AAT12221 401 bp mRNA EST 24-DEC-1997  
 LOCUS 31949 Lambda-PRL2 Arabidopsis thaliana cDNA clone 180C9T7, mRNA  
 DEFINITION  
 ACCESSION AAT12221  
 VERSION AAT12221.1 GI:2722138  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.  
 1 (bases 1 to 401)  
 Newman,T., deBrujin,F.J., Green,P., Keegstra,K., Kende,H.,  
 McIntosh,L., Ohlrogge,S., Raikhel,N., Somerville,S., Thomashow,M.,  
 Retzel,E. and Somerville,C.  
 Genes galore: a summary of methods for accessing results from  
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 On Sep 19, 1997 this sequence version replaced g1:1520373.  
 CONTACT: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313tcn@dm.cl.msu.edu  
 Seq primer: T7 dye primer.

## FEATURES

## source

1..401  
 /organism="Arabidopsis thaliana"  
 /strain="var columbiana"  
 /db\_xref="taxon:3702"  
 /clone="180C9T7"  
 /note="Vector: lambda zip-lox; Site\_1: Sal; Site\_2: Not;  
 Lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRV's lambda zip-lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dT primed cDNA." 89 t 5 others  
 BASE COUNT 104 a 75 c 66 g  
 ORIGIN











```

OY 296 cagatctagagctctacaaaacacacacacaggaactatccctcgaacttggaat 355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 537 CAGTATCTGTAGTGTACAAAACAACTCAAGAACTATTNNNNNGAACTTGGAAAC 478
OY 356 ctgaagaatctcatcatctgtgctgtacacaaacatctacaggaatgtcccaact 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 477 CTGAAGAGCTGTAGTGTGAGCTGTGACAAACAACTATCAAGCACCCTCCACT 418
OY 416 tcttttggaatctgaagctctgtctctttacggttaacacaggaattgacgag 475
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 417 TCATTGGGAAATTTGAAGAAATTTCTTTTACACATCAATGACAAACGACTATGCG 358
OY 476 ccaatccctagagcactcactcgaatcccaagccttaagttgtgtatgtctcaagaat 535
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 357 CCAATTCACAAAGAACTGTCTGTCTTCAAGCCTTAAAGTAGTAGTCTCCAACT 298
OY 536 gatttggtggaacatcccaacaaaggaactttgtctacattccttcaaggaactt 595
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 297 GATTATGTGTACAAATCTCTACATCTGGCCATTCGACATATTCATTGAATTAATTT 238
OY 596 gagaacaacccgaggttggaaggaacgaatactcgttcttcaagctacgaactaac 655
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 237 GGAATTAATCCCGCTTGGAGAGTCCAGAGTGTGGGACTAGTAAGTTATGACACAAC 178
OY 656 tgcacctgaaaaaatg 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 TGCTGCTGAATAACGTG 161

```

```

RESULT 14
AM21278 511 bp mRNA EST 07-DEC-1999
LOCUS EST297747 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION CDNA clone cLEF13N13, mRNA sequence.
ACCESSION AM21278
VERSION AM21278.1 GI:6532962
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 511)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Koning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
On Jul 7, 1999 this sequence version replaced gi:5408871.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
1..511
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_1db="cLEF13N13"
/clone_1lb="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in

```

half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

Query Match 37.6%; Score 296.8; DB 74; Length 511;  
Best Local Similarity 75.2%; Pred. No. 1.7e-71;  
Matches 370; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

```

OY 80 gtgcgaagcaactccgaagagatgctctttagcgtcttcgacgagttacagatccg 139
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 20 GTAAGGGAAATTCAGAAAGGATGCTTTGACGCCCTCCGGAGCTTATCGAACCG 79
OY 140 gaccatgtctccagaagctggaatcccaactctgttaactccttgaactgttccatgtc 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 GGTACGCTGTACAGAGCTGGATCCAAATCTTGAACCTTGTACCTGGTTTCATGTC 139
OY 200 acctgtaaccaagacacccgcgtcaactgtgtgatttgggaattcaactctcgtga 259
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 ACTTGCACAGGACATATCAATCACTGCTGTGATCTTGGAACTCAAAATTAATCTGT 199
OY 260 catcttgcgctgaagcttgggaacttgaacattacagatctagagctctacaaaac 319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CATTGTGTAAGCTGAGCTCGGAAAGCTTGAACATCTACAGTATCTGAGCTTACAAAAT 259
OY 320 aacatccaaggaactaacctccgaacttgaatctgaagaatctcatcagcttgat 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 260 AATATTCAGGGAACCATCCCTAAGAGCTGGTGAATCTGAAGAGCTTATTAATGTGAT 319
OY 380 ctgtacaacacacatcttcaaggaatagttcccaactcttcttggaaaattgaagctctg 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 320 CTGTACACAAACATATTTCCGGGACATCTCTACTTACTTGGAACCTGAAAACCTT 379
OY 440 gctctttacggttaacgaacgcatgacggggcaatccctagagcactcactcga 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 380 GTTTCTTGGCTTAATGATATCAAGCTAACAGGACCAATCCCAAGAGAACTTACTAGC 439
OY 500 atcccaagccttaaatgttgaatgttgaatgaagcaatgttcttgggaacatcccaac 559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 440 ATTTCTAGCTGAAAGTGTGGATGTCTGAAATGATGATTTGTGTGAAACATTTCTACT 499
OY 560 aacggaacttt 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 500 TCTGTCTCATTT 511

```

```

RESULT 15
AM030188 714 bp mRNA EST 15-SEP-1999
LOCUS EST273443 tomato callus, TAMU Lycopersicon esculentum cDNA clone
DEFINITION cLEC21C20, mRNA sequence.
ACCESSION AM030188
VERSION AM030188.1 GI:5888944
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 714)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Viston,T.,
Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
Koning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
On Dec 20, 1995 this sequence version replaced gi:1130947.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University

```

```

FEATURES
Source
1..511
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone_1db="cLEF13N13"
/clone_1lb="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in

```

```

TITLE
JOURNAL
COMMENT
On Dec 20, 1995 this sequence version replaced gi:1130947.
Contact: David Frisch
Clemson University Genomics Institute
Clemson University

```











GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:21:39 ; Search time 12463.9 Seconds  
(without alignments)  
-69.776 Million cell updates/sec

Title: US-09-180-798-28  
Perfect score: 894  
Sequence: 1 ggaacgatcaagcctcga.....gtaactgattcttattta 894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_om:\*  
4: gb\_ov:\*  
5: gb\_pat:\*  
6: gb\_ph:\*  
7: gb\_pl1:\*  
8: gb\_pl2:\*  
9: gb\_pr1:\*  
10: gb\_pr2:\*  
11: gb\_pr3:\*  
12: gb\_ro:\*  
13: gb\_sts:\*  
14: gb\_sy:\*  
15: gb\_un:\*  
16: gb\_yt:\*  
17: em\_fun:\*  
18: em\_hum1:\*  
19: em\_hum2:\*  
20: em\_in:\*  
21: em\_om:\*  
22: em\_or:\*  
23: em\_ov:\*  
24: em\_pat:\*  
25: em\_ph:\*  
26: em\_pl:\*  
27: em\_ro:\*  
28: em\_sts:\*  
29: em\_sy:\*  
30: em\_un:\*  
31: em\_yt:\*  
32: gb\_htg1:\*  
33: gb\_htg2:\*  
34: gb\_in1:\*  
35: gb\_in2:\*  
36: em\_ba1:\*  
37: em\_ba2:\*  
38: em\_hum3:\*  
39: em\_hum4:\*  
40: gb\_pr4:\*  
41: gb\_htg3:\*  
42: gb\_htg4:\*  
43: gb\_htg5:\*  
44: gb\_htg6:\*

45: gb\_htg7:\*  
46: em\_htg1:\*  
47: em\_htg2:\*  
48: em\_htg3:\*  
49: em\_hum5:\*  
50: gb\_pl3:\*  
51: gb\_pr5:\*  
52: gb\_htg8:\*  
53: gb\_htg9:\*  
54: gb\_htg10:\*  
55: gb\_htg11:\*  
56: gb\_htg12:\*  
57: gb\_htg13:\*  
58: gb\_htg14:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	894	100.0	894	5	A67823	A67823 Sequence 28
2	361	40.4	789	5	A67821	A67821 Sequence 26
3	361	40.4	1063	5	A67825	A67825 Sequence 30
4	356.2	39.8	981	5	A67819	A67819 Sequence 24
5	330	106	1106	5	A67817	A67817 Sequence 22
6	119.4	13.4	4604	7	LEURENE	X95269 L.esculentu
7	106	11.9	2089	5	A67827	A67827 Sequence 32
8	95.4	10.7	1755	8	DCU93048	U93048 Daucus caro
9	95.4	10.7	1815	5	A67797	A67797 Sequence 2
10	93	10.4	936	8	SBU62379	U62279 Sorghum bic
11	79	8.8	4327	32	T5F17	AL04917 Arabidops
12	77.8	8.7	2786	7	ABO29327	ABO29327 Nicotiana
13	76	8.5	14248	7	AP000815	AP000815 Oryza sat
14	73.8	8.3	87434	7	ABO10698	ABO10698 Arabidops
15	73.4	8.2	3176	5	E12705	E12705 Arabidops
16	73.4	8.2	3176	8	ATG47029	U47029 Arabidops
17	71.6	8.0	4081	5	A67815	A67815 Sequence 20
18	71.6	8.0	98471	8	ACO12654	ACO12654 Arabidops
19	70	7.8	75050	8	ACO02334	ACO02334 Arabidops
20	69.6	7.8	100887	50	ATAC011620	AC011620 Arabidops
21	68.6	7.7	111945	8	ATF1C12	AL022224 Arabidops
22	68.2	7.6	52717	7	ABO19227	ABO19227 Arabidops
23	68	7.6	96475	7	ATF17M5	AL035678 Arabidops
24	67.6	7.6	81513	50	ACO21198	ACO21198 Arabidops
25	67.6	7.5	108355	50	ACO05957	ACO05957 Arabidops
26	67	7.5	143186	8	ACO05287	ACO05287 Arabidops
27	66.4	7.4	3573	5	A57133	A57133 Sequence 4
28	66.4	7.4	3568	8	U42445	U42445 Lycopersico
29	66.4	7.4	6471	5	A57130	A57130 Sequence 1
30	66.4	7.4	6471	8	U42444	U42444 Lycopersico
31	66.4	7.4	83649	50	ACO05395	ACO05395 Arabidops
32	66.2	7.4	116944	8	ACO08017	ACO08017 Arabidops
33	65.8	7.4	99725	8	ATF18F4	AL021637 Arabidops
34	65.6	7.3	2936	8	AF053994	AF053994 Lycopersi
35	65.2	7.3	65899	50	ACO06436	ACO06436 Arabidops
36	64.8	7.2	94362	8	ACO05405	ACO05405 BAC F18A1
37	64.8	7.2	127053	8	T1N24	AF149413 Arabidops
38	64.6	7.2	75568	50	ACO06532	ACO06532 Arabidops
39	64	7.2	97711	8	AT16K5	AL132965 Arabidops
40	64	7.2	104204	8	AT19C5	AL132964 Arabidops
41	63.4	7.1	114935	8	ACO07651	ACO07651 Arabidops
42	63	7.0	3541	5	A67432	A67432 Sequence 5
43	63	7.0	3541	8	AF053397	AF053397 Lycopersi
44	63	7.0	6695	5	A67796	A67796 Sequence 1
45	62.8	7.0	3979	5	A67428	A67428 Sequence 1

## ALIGNMENTS



[illegible]

QY	541	tcagggaatgatttctgtggaacaattccgtgaagagccctttbaacacttctatg	600
QY	541	tcagggaatgatttctgtggaacaattccgtgaagagccctttbaacacttctatg	600
Db	541	TCAGGGAATGATTTGTGTGGAACATTCGATGAGAGACCTTTTGAACACATTCCTATG	600
QY	601	caaaactctgagaacaccctgagatctgagagacagaaactaagcttctgcagctat	660
Db	601	CAAAACTCTTGAGAACACACTGATGTGAGGAGCACCAACTACTAGCTCTTCGAGCTAT	660
QY	661	gacaccaattgcacctaataaagaagtctgaagaaccttaagaagaatttctgtgact	720
Db	661	GACACCAATTGCACTTAAAAAGAACTTGAACACCTTTAAAGAGATGTATAGTGACCT	720
QY	721	tgtaagaactctgtaccaaagtcttctgaatacctatataagacccttcttcaatgata	780
Db	721	TGTAAAGACTCTGTACCAAGTCTTTGTAATACCTATATAGAGCTTGTTCATGTTATATA	780
QY	781	tgaagcttggagaagacagtaactctgcaatgtatgttattgtttaaaaaagttaagt	840
Db	781	TGAAGCTTTGAGAAGACACTACTTGCATATTTGTTATGTGTGTAAGAAAAGTTGAATG	840
QY	841	agaattgcttgaattgattgttcttcttaactgaactgaattcttaata	894
Db	841	AGATTGCTTTGTAATGTGATTTGTCTTTATGTAACCTGGAATTTCTATATA	894
RESULT	2		
LOCUS	A67821	789 bp	DNA
DEFINITION	Sequence 26 from Patent WO9743427.		PAT
ACCESSION	A67821		
VERSION	A67821.1	GI:4756643	
KEYWORDS	unidentified.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 789)		
AUTHORS	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.		
TITLE	PRODUCTION OF APOMICRIT SEED		
JOURNAL	Patent: WO 9743427-A 20-NOV-1997;		
FEATURES	CTBA GEIGY AG (CH)		
source	Location/Qualifiers		
	1..789		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
	<2..664		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CBA42248.1"		
	/db_xref="GI:4756644"		
	/translation="RPTPRRYRWELFPAASLIITLALHLVANESEGAIALALRSIT		
	DPHYVQSWDPPLVNPCTWPFHVCNDONRYTRVDLGNLSLSHLAPELGKLEHLQYLE		
	LXNNINQGIIPSELGNLKNLISDLVNNNTLTGIVPSLGKLSLVEFLNDNRLLTGPI		
	PRALTRIPSLKIVDVSSNDLCGTIPITNGPFAHIPIQNFENNRLTEGPILLGLASYDTNN		
CDS			
BASE COUNT	235 a	191 c	156 g
ORIGIN			207 t
Query Match	40.4%;	Score 361;	DB 5;
Best Local Similarity	72.8%;	Pred. No. 6.9e-73;	Length 789;
Matches 538; Conservative	0;	Mismatches 110;	Indels 91;
		Gaps	2;
QY	15	ctccgaagggaagcctcttcacgcgcttcgcggagctatcacgaatcagaatgtgt	74
Db	91	CTCCGAAGGAAGATCCTCTTACCTCTTCCGCGGAGTTAAACAATCCGACATGCTCT	150
QY	75	tcaagattggatacaactcttgtaactctgtacttggttcaatgtaataca	134
Db	151	CCAGAGCGTGGAGTCCAACTCTTGTAACTCTGTACCTGATGCTCATGACACTGTAAACA	210
QY	135	acacataaagtcacgtctgtgatttgggaatcaactatctggacatagtaacc	194
Db	135	ACACATAAAGTCACGTCGTGATTGGGAATCAACTATCTGGACATAGTAGTAACT	194



Db 211 AGACAAACCGGTCACCTGCTGTGATTTGGGGAATTCAAACCTCTGTGACATCTTGCCGC 270  
 QY 195 tgaacttgggaagcttgaacatttacaatcatgtatgaatcatcactctttgcttt 254  
 Db 271 TGAGCTTGGGAAGCTTGAACATTTACAGTATCT----- 304  
 QY 255 tgattatcgtaaaaacattacattatcagtcacacataataacatttgcttgatcata 314  
 Db 304 ----- 304  
 QY 315 tagtgaacttacaacacagagattcaaggaactatccttgaagcttgaacattgaa 374  
 Db 304 ---AGAGCTTACAAAACAACTCCAGGACATACCTTCCGAACTTGGAAATCTGAA 360  
 QY 375 gagtctaactgaatttgatctgtacacaaacattcaccggyaaaatccactcttt 434  
 Db 361 GAATCTCATACGCTTGATCTGTACACAACATCTTACAGGGAATAGTCCACATCTCTT 420  
 QY 435 gggaaaattgaagctactgttttttttggggttaacgaacacgattgacggtccat 494  
 Db 421 GGGAAATTTAAGTCTCTGTCTTTTACGGCTTAATGACAAACGATTGACGGGCCAAT 480  
 QY 495 tcttagagaactcacaagtatttcaagccttaagttgttgaatgtctcaggyaatgatt 554  
 Db 481 CCCTAGAGCACTCACTGCAATCCCAAGCCTTAAGTTGATGCTCAAGCAATGATTT 540  
 QY 555 gtgtggaacattccagtagaagacctttgaacacattcctatgacaaacttggaa 614  
 Db 541 GTGTGGAACATATCCCAACAAACGACCTTTGTCTACATTCCTTACAAACTTTGAGAA 600  
 QY 615 caacttgaatttggaggaacacaaactaagtgcttggagctatgaacacaaattgac 674  
 Db 601 CAACCCGAGGTGGAGGAGCAGGAATTAAGTCTTGCAGAGTACGACATCACTGAC 660  
 QY 675 ttaaaaagaagtggaagaacctataagaaga-tgttggtggaacttgaagactctg 733  
 Db 661 CTGAAAAAATTTGGCAAACTGTAAGAAATGAGAAATTTGGGGGTGACCTGTGAAGACACTT 720  
 QY 734 taccgaagtgttgaatc 752  
 Db 721 CACCACCTTTATCAAAATATC 739  
 RESULT 3  
 A67825 1063 bp DNA PAT 05-MAY-1999  
 LOCUS Sequence 30 from Patent WO9743427.  
 DEFINITION A67825  
 ACCESSION A67825  
 VERSION A67825.1 GI:4756647  
 KEYWORDS  
 SOURCE unidentified.  
 ORGANISM unidentified.  
 REFERENCE 1 (bases 1 to 1063)  
 AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
 TITLE PRODUCTION OF APOMICRITIC SEED  
 JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
 CITERA GEICY AG (CH)  
 FEATURES  
 source Location/Qualifiers  
 1..1063  
 /organism="unidentified"  
 /db\_xref="taxon:32644"  
 106..762  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAB42252.1"  
 /db\_xref="GI:4756648"  
 /translation="MASRNYRWELFAASLITLALIHVANSSEGDALYALRSITDP  
 DHVLSMDPLVPCFTWTFVTCNDNRVTVLDGNSNSGHLAPGLGHEHQYLEIK  
 KNNIOGTIPSELGNLKNLSLDLYNNNLGIYVTSIGLKSIVFLRLDNRITGP  
 ALTAIPSLKVVYDSSNDLGGITPTNPFPHIPLQNFENNPRLEGEELGLASIDTNCI  
 "

ORIGIN  
 Query Match 40.4%; Score 361; DB 5; Length 1063;  
 Best local similarity 72.8%; Pred. No. 6,7e-73;  
 Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2;  
 QY 15 ctccgaagggaagcgtctcagcgcgttcgcgcggaactatcagatccagaatgtgt 74  
 Db 189 CTCGGAAGAGATGCTCTTACGCTCTTCGCGGAGTTAAACATCCGACATGTTCT 248  
 QY 75 tcsagttggatccaaactctgttaactccttgatctgttcaacttcaacttaata 134  
 Db 249 CCAGAGCTGGAGATCCACTCTGTAACTGTAACTGTGACGTGACATGACCTGTAACA 308  
 QY 135 acacacataagtcactcgtcgttgatttgggaattcaaacattatcagactagatcc 194  
 Db 309 AGACAACCGGTCACCTCGTGTGATTTGGGGAATTCAAACCTCTGTGACATCTTGCGCC 368  
 QY 195 tgaacttgggaagcttgaacattacatatctgtatgtgaatcatcactcttgcctt 254  
 Db 369 TGAGCTTGGGAACCTTGAACATTTACAGTATCT----- 402  
 QY 255 tgattatcgtaaaaacattacattatcagtcacacataataacatttgccttgatcata 314  
 Db 402 ----- 402  
 QY 315 tagtgaacttacaacacagagattcaaggaactatcactctgaagcttgaacattgaa 374  
 Db 402 ---AGAGCTTACAAAACAACTCCAGGACATACCTTCCGAACTTGGAAATCTGAA 458  
 QY 375 gagtctaactgaatttgatctgttacaacacaaactcaccggaacatccactcttt 434  
 Db 459 GAATCTCATACGCTTGATCTGTACACAACATTTACAGGGAATAGTCCACATCTCTT 518  
 QY 435 gggaaaattgaagctactgttttttggggttaacgaacacgattgacggtccat 494  
 Db 519 GGGAAAAATGAAATCTCTGTCTTTTACGGCTTAATGACAAACGATTTGAGGGGCCAAT 578  
 QY 495 tcttagagaactcacaagttatttcaagccttaagttgttgaatgtcgaaggaaatgt 554  
 Db 579 CCTTAGAGCACTACCTGCAATCCCAAGCCTTAAGTTGATGCTGCTACAAATGATTT 638  
 QY 555 gtgtggaacaaatccagtcgaaggaaccttgaacacattcctatgacaaacttgaaga 614  
 Db 639 GTGTGGAACATCCCAACAAACGACCTTTGTCTGCAATTCCTTACAGAACTTGAANA 698  
 QY 615 caactggaatttggaggaacccgaactaagtgcttgcgagctatgacaccaattgac 674  
 Db 699 CAACCCGAGGTGGAGGAGCAGGAATTAAGTCTTGCAGACTGACACTTAAGTGCAC 758  
 QY 675 ttaaaaagaagttgaagaaacctataagaaga-tgttgaagtcacttgaagactctg 733  
 Db 759 CTGAAAAAATTTGGCAAACTGTAAGAAATTTGGGGGTGACCTGTGAAGAACTT 818  
 QY 734 taccgaagtgttgaatc 752  
 Db 819 CACCACCTTTATCAAAATATC 837  
 RESULT 4  
 A67819 981 bp DNA PAT 05-MAY-1999  
 LOCUS Sequence 24 from Patent WO9743427.  
 DEFINITION A67819  
 ACCESSION A67819  
 VERSION A67819.1 GI:4756641  
 KEYWORDS  
 SOURCE unidentified.  
 ORGANISM unidentified.  
 REFERENCE 1 (bases 1 to 981)  
 AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
 TITLE PRODUCTION OF APOMICRITIC SEED



JOURNAL		Patent: NO 9743427-A-20-NOV-1997:	
FEATURES		CIBA GEIGY AG (CH)	
source	Location/Qualifiers		
CDS	1..981 /organism="unidentified" /db_xref="taxon:32644" 104..760 /note="unnamed protein product"		
BASE COUNT	286 a 236 c 180 g 279 t		
ORIGIN			
Query Match	39.8%; Score 356.2; DB 5; Length 981;		
Best Local Similarity	72.4%; Pred. No. 8.5e-72;		
Matches	535; Conservative 0; Mismatches 113; Indels 91; Gaps 2;		
QY	15 ctcggaaggagacgcctctacgcgccttcgcggagcttatagatccagacaatggtt 74		
DB	187 CTCGCCAAGAGAGATGCTCTCTACGCTCTTCGCCGGAGTTGACAGATCCAGACATGCTCT 246		
QY	75 taaggatggatccaaacctctgtaacctctgtactgttcttgatctgtaactgtatca 134		
DB	247 CCAGAGACTGGATTCACACTCTGTTATCTCTGTACTCTGTTCATGTCACTGTACCA 306		
QY	135 acacacatcaagctactcgtctgtgatttggggaattcaaaactatctgacatctagacc 194		
DB	307 AGACAAACCGCGTACACCTCGTGTGGAATTCAAACCTCTCGACATCTTGCGCC 366		
QY	195 tgaacttggaaagcttggaaacattcaaatctgtctgtaacatcatcactcctttgcctt 254		
DB	367 TGAGCTTGGGAAGCTTGAACATTTAAGATCT----- 400		
QY	255 tgattatctgaanaacattacattacagctcaacacataataattgtcttgagtcata 314		
DB	400 ----- 400		
QY	315 tagtgaactctcaaaaacagagattcaaggaactatacctctgagcttggaaatctgaa 374		
DB	400 ---AGAGCTCTACAAAAMACATCCAAAGAACTATACCTCCGAACCTGGAAATGTGAA 456		
QY	375 gagctcaatcagatttggatctgtacaaacaacaatctccacggggaataatccactctctt 434		
DB	457 GAATCTCATCAGCTTGGATCTGTACAAACAACATCTTAAAGGATATGTTCCACTTCTT 516		
QY	435 ggggaaatgaaagtcacttgttttttttgcgcttaagaaacagattgacgcgtctat 494		
DB	517 GGGAAATTTGAAGTCTCTGCTTTTAAAGGCTTATGCAACCGGATTAACGCGTCAAT 576		
QY	495 tcttagaagactcaacgattatccaagcctttaaagtgttgatgtrctggggaattat 554		
DB	577 CCTTAAGACCTACGCGCAATCCAAAGCCTTAAAGTGTGAGGTCAAGCATATATT 636		
QY	555 gttgtgaacaatctcagtagaagacacctttgaaacatctctatgcaaaactttgaga 614		
DB	637 GGTGGAACAATCCCAAAAGGAGCCCTTGTGCATTCCTTTACAGATCTTGAGAA 696		
QY	615 caactcgagatttggaggagcagaactactagtcttgcagctatgacccaattggac 674		
DB	697 CAACCCGAGATTGGAGGAGCCGGAATTAATCTCGCTTTGCAAGCTAGACACTAACGCA 756		
QY	675 tttaagaagaagtggaaagacctataagaagaa-tgttggtgtaaccttgaagaactctg 733		
DB	757 CTGAACAACACTGGCAAAACCTGAATAATGAAGATTTGGGGGTGACTTTGAACAACATT 816		
QY	734 taccaagtglttgtaaatc 752		

Db	817	CACGACCTTATCAAAATATC	835	
RESULT	5			
LOCUS	A67817	1106 bp	DNA	PAT
DEFINITION	Sequence 22 from Patent WO9743427.			05-MAY-1999
ACCESSION	A67817			
VERSION	A67817.1	GI:4756639		
KEYWORDS				
SOURCE	unidentified.			
ORGANISM	unidentified.			
REFERENCE	unclassified.			
AUTHORS	1 (bases 1 to 1106)			
TITLE	De V.S., Schmidt, E.D., Van, H. G. and Hecht, V.F.			
JOURNAL	PRODUCTION OF AROMATIC SED			
	Patent: WO 9743427-A 20-NOV-1997;			
	CIBA GEIGY AG (CH)			
FEATURES	Location/Qualifiers			
source	1..1106			
	/organism="unidentified"			
	/db_xref="taxon:32644"			
CDS	142..798			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAB42244.1"			
	/db_xref="GI:4756640"			
	/translation="MASRNYRWELEFAASLTETLALIHVEANSSEGLALYLALRSLRDP			
	DHVALSDPILVNCPTFWHTGQNDNNRYTVDLNSLSGLHAPLEKLEHIOYLELY			
	KNNIGQITPSLEKLNKLISIDLYKNNLTGIVPEFLKLSIVPLRLNDRINQPILE			
	HSRSQAFKVVVDSNDLCGITPNGFAHLPFLQNFENRPLRLESPLELGLASIDTNC			
BASE COUNT	331 a	258 c	206 g	311 t
ORIGIN				
Query Match	36.9%; Score 330; DB 5; Length 1106;			
Best Local Similarity	71.9%; Pred. No. 8.3e-66;			
Matches 532; Conservative	0; Mismatches 115; Indels 93; Gaps 4;			
QY	15	ctccgaaggagagcgtctcttcacgcgcttcgcgcgagactatcagaatcgaagaatgctgt	74	
DB	225	ctccgaaggagagatgctctcttctttagcgtcttgcgcgagttggaagatccagaccatgctct	284	
QY	75	tcagagttggagatcgaactctgtttaactctgttaactctgttctcaatctgtgaatc	134	
DB	285	ccagagctggagatcgaactctgtttaactctgttgaactctgttgcacatgctgaaccca	344	
QY	135	acacatcaatgaactcgtctgtgatttggagatcaaatcattatctggaatctagtaacc	194	
DB	345	agacaaacccgcgcacactcgttgatttggagatcaaatcatttggacatcttgccgc	404	
QY	195	tgaacttggagagcttgaacatttcaatcattatctgtaatgaatcactctttgcctt	254	
DB	405	tgaacttggagagcttgaacatttcaatcatttgaatgaatcactctttgcctt	438	
QY	255	tgattatctgaaacatttcaatcattatctgtaatgaatcactctttgcctt	314	
DB	438	-----	438	
QY	315	tagtgaactctcaaaaacgagatcaagaactatactcttgagcttgaatctgaa	374	
DB	438	-----	438	
QY	375	gagctcaatcaagtttgatctgtacaacaaacatctcaccgaggaatccatctctt	434	
DB	495	gaatcctacacacttgatctgtacaaacatcttgaaggaatggtccacatttctt	554	
QY	435	gggaataatgaagctcgttttttttgggttaagaaacgattgacagctccat	494	
DB	555	gggaataatgaagctcgttttttttgaaggttaatgaacacgattgacagctccat	613	



[illegible]







DEFINITION Sequence 2 from Patent WO9743427.  
ACCESSION A67797  
VERSION A67797.1 GI:4756623  
KEYWORDS  
SOURCE carrot.  
ORGANISM Daucus carota  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.  
REFERENCE 1 (bases 1 to 1815)  
AUTHORS De V.S., Schmidt, E.D., Van, H.G. and Hecht, V.F.  
TITLE PRODUCTION OF APOMITIC SEED  
JOURNAL Patent: WO 97/43427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
source Location/Qualifiers  
1..1815  
/organism="Daucus carota"  
/db\_xref="taxon:4039"  
94..1755  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42241.1"  
/db\_xref="gi:4756624"  
/translation="MNRSINILNTMOTFDAYLDKYGVLNLTLEYNNISGPIPSDLG  
NLTVSLDILVNSSPGIPDILGKLRRLRLNLSGPIPSMSTLITLOVLD  
SNRLSGVPNGSFLPTISPANNINLGPVGRPCSPSPSPPTIPSTOP  
PGONGPTEAINGVYAGACALFPAPAPAFMRPRRPFVRAEDPEVILGOLK  
RRLSELVATDITSTILGRGFGKVKGRADSLVAVRKLKERTPGDELQTEY  
EMISAVRNLRNGFCMTPELRLVYPIVANGSVASCLEROPSPLDMPTRRI  
ALGSAKSLYLDHDCDPIIRDVKANILDEFEAVVDGFLARLMDKDHVTA  
VAGTGLYLAPELSTGKSEKTDVGVGIMLELITGORAFLARLMDKDHVTA  
KSLLEKLEMLVDPLENNYIDTEVEVDLVALCTGSGSPMERPKSEVVRMLEDGG  
LAERDEKQEVYIHQVVELAPHRSEWILDSTONLHAFELSGPR"  
BASE COUNT 531 a 354 c 415 g 515 t  
ORIGIN  
Query Match 10.7%; Score 95.4; DB 5; Length 1815;  
Best Local Similarity 56.1%; Pred. No. 3.7e-12;  
Matches 180; Conservative 0; Mismatches 141; Indels 0; Gaps 0;  
QY 319 gaacttacaacaaagattcaagaaactaaccttcctggagcttggaactgaagagt 378  
DB 178 GACCTTTACGCAATTAACATTAAGTGGACCAATTCCTAGGATCTGGGAATCTGACAAAT 237  
QY 379 ctaacagttggtatctgtacaacaacaaatctcacccgggaaatccactctcttgga 438  
DB 238 TTGGTGAAGCTTGACCTATATGATAGCTTCTGTGACCTATACCGGACAAATTAAGA 297  
QY 439 aaattgaagtcactgttttttttgggcttaacgaaacggaattgaacgggtccatttcc 498  
DB 298 AACCTTACAAGGCTTAAGATCTTGCTGCAACAAACAGAGCTCTGGTCCCAATTCA 357  
QY 499 agagaactcacagttattcaagccttaagttgttgatgtctcagggaatgaattggt 558  
DB 358 ATGCACTGACTAATATTAACAATCTTCAAGTCCGGATTTATCAACAATCGGTATACA 417  
QY 559 ggaacaattccagtagaaggaacttttgacaacattcctatgcaaaactttgagaacaac 618  
DB 418 GGACCACTACCGGATTAATGAGCTTATTTCTTTTACACCTATACGATTTTGCCAAATAT 477  
QY 619 ctgagattggaaggaccagaa 639  
DB 478 TTGAATTATATGTGACCCGTA 498  
RESULT 10  
LOCUS SB062279 936 bp mRNA PLN 09-DEC-1996  
DEFINITION Sorghum bicolor leucine-rich repeat-containing extracellular  
glycoprotein mRNA, complete cds.  
ACCESSION U62279  
VERSION U62279.1 GI:1710123

KEYWORDS sorghum.  
SOURCE Sorghum bicolor  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphylliphytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.  
REFERENCE 1 (bases 1 to 936)  
AUTHORS Hipskind, J.D., Nicholson, R.L. and Goldsborough, P.B.  
TITLE Isolation of a cDNA encoding a novel leucine-rich repeat motif from Sorghum bicolor inoculated with fungi  
JOURNAL Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)  
MEDLINE 97124217  
REMARK Erratum: [[published erratum appears in Mol Plant Microbe Interact  
1997 Mar;10(2):302]]  
REFERENCE 2 (bases 1 to 936)  
AUTHORS Hipskind, J.D.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant  
Pathology, Purdue University, West Lafayette, IN 47907, USA  
FEATURES  
source Location/Qualifiers  
1..936  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone="SLRR"  
12..735  
/function="protein-protein ligand binding"  
/note="leucine-rich repeat-containing extracellular  
glycoprotein; contains six N-glycosylation sites  
[NX(S/T)]"  
/codon\_start=1  
/protein\_id="AAC49559.1"  
/db\_xref="gi:1710124"  
/translation="MAPQFAAGFLNGILATFAACSTEGDILYKORLWEDPNVLT  
OSWSTLANCTPHVYCNNNNNVIVITVGGNGISRLIPDLAQLQVITLYNGL  
NGSIPETLGNLNLISLDLNDNLITETIPITGSSSTLRLKLRNLAGPIPSFGN  
LTSLSKSLQENSGAIPASLGNKALPRLNDMLNGTVPKSFPLTFGNLTEL  
NTDRNLDGRTSSGLNVAI10DALRTA"  
BASE COUNT 247 a 232 c 207 g 250 t  
ORIGIN  
sig\_peptide  
Query Match 10.4%; Score 93; DB 8; Length 936;  
Best Local Similarity 51.8%; Pred. No. 1.4e-11;  
Matches 250; Conservative 0; Mismatches 215; Indels 18; Gaps 1;  
QY 15 ctccgaaggggagcgtcttcacgcgcttcgcggaagcttatccagatccagaatgtgt 74  
DB 83 CACTGAAGGTGACATACGTATACAGCAAGGTTGCAATGGAGAGCCCAACACGTGCT 142  
QY 75 teagaattggatccaaactctgttaactctgtactgtgttccatgtcactgttaata 134  
DB 143 GGAAGAGCTGGAATTCGACGCTTGCAATCCTGCACTGGTTCATGTACCGCAACAA 202  
QY 135 aacaccataagtcacatgctcgtgatttgggaattcaaaacttactcgtgaacattagtaac 194  
DB 203 CAATATCTTTGATCCCGGTGGATTTGGCAATGAGCATCTCGGTCTGTGCTTC 262  
QY 195 tgaacttgggaagcttgacaattcaatatactgtatggaatacacaactctcttgccttt 254  
DB 263 AGATCTTGCAAAATTCGAACCTCCATATCATGAGGTGATGACGACGCTTGAACGG 322  
QY 255 tgaattatcgaataaatttacaattatcagttacacataataacatttgccttgaatca 314  
DB 323 TTCGATTCGAGAAACAGTAGGCAACCTGACATATCATCATGCTT----- 368  
QY 315 tagtgaacttacaacaaagagattcaaggaactataccttctgaagttggaatctgaa 374  
DB 368 ---GGATCTCTGGGCAACACTTCTTACCGGGAATCCCAACTACGCTTGTTCTGTGCA 424  
QY 375 gagttcaatcagtttggatctgtacaacaaacttcaacgggaaatcccatctcttt 434



Db 425 CACGCTGCATATCTGAGTTGTACCAAAACACCTGACAGGCGCTATACCATCATCGTT 484  
QY 435 gggaaatgaaagtcattgttttttcgagcttaagaaacgattgaccggtcctat 494  
Db 485 TGGCAACCTGACTACGCTTGTGAATCGAAGCTTCAGAGATTCGTTGAGCGCGCTAT 544  
QY 495 tcc 497  
Db 545 TCC 547

RESULT 11  
T5F17/c  
LOCUS T5F17 43277 bp DNA HMG 06-AUG-1999  
DEFINITION Arabidopsis thaliana clone T5F17, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
in unordered pieces.  
ACCESSION AL049917.1 GI:4884021  
KEYWORDS HMG; HMGs\_PHA5E1.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsi  
1 (bases 1 to 43277)  
Lenhard,N., Quail,M., Harris,B., Rajandream,M.A. and Barrell,B.G.  
Direct Submission  
Submitted (20-MAY-1999) EU Arabidopsis sequencing project. The  
Sanger Centre, Wellcome Trust Genome Camps, Hinxton, Cambridge  
CB10 1SA, UK  
COMMENT IMPORTANT: This sequence is unfinished and does not necessarily  
represent the correct sequence. Work on the sequence is in progress  
and the release of this data is based on the understanding that the  
sequence may change as work continues. The sequence may be  
contaminated with foreign sequence from E.coli, yeast, vector,  
phage etc.  
For more information about this sequence or the ESSALI project, see  
MIPS <http://webvr.mips.biochem.mpg.de/proj/thal/>. Order of  
segments is not known, 800 n's separate segments.  
\* NOTE: This is a 'working draft' sequence.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

FEATURES  
source Location/Qualifiers  
1..43277  
/organism="Arabidopsis thaliana"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/clone="T5F17"

BASE COUNT 13207 a 7883 c 8609 g 13578 t  
ORIGIN

Query Match 8.8%; Score 79; DB 32; Length 43277;  
Best Local Similarity 57.5%; Pred. No. 1.6e-08;  
Matches 142; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 329 aaaaagagattcaaggaataccttcgaacttgaactgaagcctgaagccttaacagtc 388  
Db 30596 ACAGCAAAATTCAAGAGTCCGATTCCACCGAATTGGAAACATAAAGCCTCAAGTACC 30537  
QY 389 tggatctgtacaacaacaatctaccgaggaataatccatctcttcttggaaatgaagt 448  
Db 30536 TTGATTTGGCCATCGAAGGCTCAGTGGAGATTCCCTCAGAGCTTGGGAAGCTCAAGT 30477  
QY 449 cactcttttttttggagcttaacgaaacgattgacggctcctattcctgaagaactca 508  
Db 30476 CACTCGAGACGCTTCTCTGTACGAGAACTTCAACCGGAACATTCCAGAGAGTAG 30417  
QY 509 cagtaattcaagccttaagttgtatgctcaggggaatgatttgttggaaacttc 568  
Db 30416 GAAGCAATTACACAGCTAAGGTTCTGATTCTCCGATAATGCTTAACCGGAGATTTC 30357

QY 569 cagrtaga 575  
Db 30356 CAATGGA 30350

RESULT 12  
AB029327  
LOCUS AB029327 2786 bp mRNA PLN 25-DEC-1999  
DEFINITION Nicotiana tabacum mRNA for elicitor-inducible receptor-like protein  
EIR, complete cds.  
ACCESSION AB029327  
VERSION AB029327.1 GI:6635235  
KEYWORDS elicitor-inducible receptor-like protein EIR.  
SOURCE Nicotiana tabacum mRNA.  
ORGANISM Nicotiana tabacum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
Nicotiana.  
1 (sites)  
Takemoto,D. and Kawakita,K.  
Isolation of the gene for EIR, an elicitor-inducible receptor-like  
protein, from tobacco by Differential Display  
Unpublished (1999)  
2 (bases 1 to 2786)  
Takemoto,D. and Kawakita,K.  
Direct Submission  
Submitted (24-JUN-1999) to the DDBJ/EMBL/Genbank databases. Daigo  
Takemoto, Nagoya University, Plant Pathology Laboratory, School of  
Bioagricultural Sciences, Chikusa, Nagoya, Aichi 464-8601, Japan  
(E-mail:1972006d@mbn.med.nagoya-u.ac.jp,  
Tel:81-52-789-4031(ex.4031), Fax:81-52-789-5525)

FEATURES  
source Location/Qualifiers  
1..2786  
/organism="Nicotiana tabacum"  
/db\_xref="taxon:4097"  
1..2586  
/gene="elir"  
1..2586  
/gene="elir"  
/codon\_start=1  
/product="elicitor-inducible receptor-like protein EIR"  
/protein\_id="BA88636.1"  
/db\_xref="GI:6635235"  
/translation="MMWYGRFLSLHFFTLFCETVYFASRTKREATALMKRAITLQNS  
NELYSMPSSKACKSWCYCVRNGRSKIDIPAGYIGTLNPPSSLPLEYIDIS  
MQULRSIPPEIGKLTNLYIDLFSPNLSGTIPPOISLAKLTTHILNHLNGSIPG  
EIGHLRSLELDLSTNLNGSIPSPDLNHLNLCLYKNNISGFIPSEIGYSSLIQ  
LDLNTNFNGSIPASLENLNLNLSTLYENQLSGSDPEIDGRLTLDRLNTNFNG  
SIPASIGNTLSLSIQLEHNLGSDIPEEIGYKTLVLSLYNPNFNGSIPISGLMT  
SLSLSLYENHLSGPIPSLGNLNLVLYLYANQSGPIPSGLMKLNMYKMLHDN  
OLNGSIPASFGNLRMNYLPLESNNLNGEIPLSICNLMSLKVSLGRNSLKGPILOCL  
INISRYOYKTRPDNLSREIPSSICNLSTRIIDLSNNKSGSIPOCFDMSGHLEVL  
DIHKNISGSLTPTTRISYVRSFTLHNELECKTIPSLANCKEIDYLDIGNLDLT  
FPMWGLTEPKDLNLRKSNKLKGSIRKSKDENMFLERITNLSTNAPFGIPSLDQ  
LKAMRKIDQVYVEPTLYLGRFGADIRENYSTVYTKLELKLIVRIITVYIITDLSNR  
FGHVPISIGELIALRVNLNRNGIOGHIPSLGNLNLVLSIDLSFNOJSGEIPQIA  
SOLSTSLVNLSTLYNLOGCIPQGFQPFHFNENSSEGDGRGPIPKSGCGNDVSETN  
NIVSTLDDQESSEFLNDFWKAALMGSGSICISLILYMISTGKLMVSRTEWLD  
NRFIT"

BASE COUNT 818 a 520 c 502 g 946 t  
ORIGIN

Query Match 8.7%; Score 77.8; DB 7; Length 2786;  
Best Local Similarity 57.0%; Pred. No. 3.8e-08;  
Matches 142; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 323 tctacaaaagagattcaaggaataccttcgaacttggaaatcgaagccttaac 382  
Db 950 TGTACTACTACTTCTTAATGTTCCATTTCCATTGCGGAATTTGACGAGCTTGT 1009







CDS  
complement(join(34468..34544,37202..37385))  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAAB7827.1"  
/db\_xref="GI:6498424"  
/translation="MAADKAGGDDAAAAAPLVSAAGRRRCPGCTLTKCKADA  
GIPYLFNFYIWWVCISFPFRPAKKICDSIVKAMAGHRR"  
join(40347..40358,40468..40680)  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAAB7828.1"  
/db\_xref="GI:6498425"  
/translation="MHPKNNVAAATVILVQSLPSSFSFNRRRRRQCPTAIAT  
ADGSTRVLPDPDGGKGFRRMRKKEGGD"  
complement(join(42025..42083,43481..43661))  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAAB7829.1"  
/db\_xref="GI:6498426"  
/translation="MAAGKAGGDDAAAAAPLVPAAGRRRCPCGCTKRCENHPG  
IPYLFNFYIWWVCISASAVNRVETDEAPPTSVS"  
join(47081..47084,47194..47313,47469..47584)  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAAB7830.1"  
/db\_xref="GI:6498427"  
/translation="MSRGRAPATACGSRVCAEVAPEQDSARYALLLRDLSP  
AATPAKLRIRHDDGVPLERDDELVSQPTD"  
join(48412..48822,49072..49179,49283..49438,49876..50091,  
50202..50411,50488..50625,50842..50943,51105..51170,  
51274..51358,51454..51551)  
/note="Similar to Arabidopsis thaliana chromosome II BAC  
17A16 sequence; hypothetical protein. (AC005496)"  
/codon\_start=1  
/protein\_id="BAAB7831.1"  
/db\_xref="GI:6498428"  
/translation="MYLILPISVCVSPAPSRRLDPDGGRGVDLGGVRF  
KRAALGVIRSWTSSSSRKGRHDDARDIGYDGDGAGALLAYRLLSGAQD  
AAEGEEDEQGFPRMAIVFLCFSAFLICMMDRNMSIALTPSAEAGNPQVLG  
IOSGFPMGILTOIAGIADVAVGKTIVFEGVIMWSIATLTPPAKIGLPIFLYR  
AFEGVEGAAAMPANNILSKWVPSESRSLALVSGMYLGSVGLASPLILNFGW  
PSFVSEGSIVGFWEFTWASKVSSPLEDPGISAEKKLITQTTGGEPRKEIPGLI  
LSKRPWALIVSHCHNMGFTFILLTMPYVNOVKFNPLETSGICVLPMLMAVSAN  
FGWADILVSRSLSTTYTRAKINOSIGFPGAPFLQLSHDISPMAYICACSGCTD  
AQSQSLSHQDIPRIVAGVLGLSNTAGVAGFTATGTIIQHGSMQDVEFVSV  
YLVLTGLWNLFSTEKIID"  
complement(join(52616..52685,53228..53299,53395..53466,  
53657..53800,54663..54734,56026..56234))  
/note="ESTs C97644(C60871), A0081265(C60871) correspond to  
a region of the predicted gene; similar to L.esculentum  
LPR gene. (X95269)"  
/codon\_start=1  
/protein\_id="BAAB7832.1"  
/db\_xref="GI:6498429"  
/translation="MAAASPALAAVLAAVASANSSEGALYLRALADPGVLIQ  
SMQPLTNPCTEHTVCDRAGRVTRLDIGLSNLSGLADELGELEHOLEYLNKNIQ  
GTPALGSLKNIISLDIVNNNTGTIPEIGLSTSLVLRINDSNLNPTRDLAKI  
SSKALVDSNNDLCGTTIPSGPEHPIPLNFDKPLDEPQLGATVYTYNC"  
53608..53826  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAAB7833.1"  
/db\_xref="GI:6498430"  
/translation="MKLIKDHDRSHHTVLYLASLSTALYTSLTIYKGTGLAQPGE  
KTGTSILGVSPEKDSRRKRPVGNLADH"  
complement(join(63689..63842,63844..64145))  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAAB7834.1"  
/db\_xref="GI:6498431"  
/translation="MRSRKATDLRAAGRAATLRRGRGSRGQGEVYLGASLGGG  
EGSDSWBERGGGGAAPAGGERRRRHAEKTRGLGSLIVGGRGAPPHRSGGSGH  
EGGDEGDELWGLGALRPEEEEGGRRRRGRGERRRVCREEER"

	Query Match	8.5%;	Score 76;	DB 7;	Length 142418;				
	Best Local Similarity	65.1%;	Pred. No. 7,3e-08;						
	Matches 112;	Conservative	0;	Mismatches	60;	Indels	0;	Gaps	0;
Qy	298	tttgtcttgatcatalagtaagtacctaatacaaacgagaattcaaggaaactataacctct	357						
Db	53820	TTTTTGAAAAATTGTATACAGAGAGCTGACAAAACAATTTTCAGGGCAGCTATCCAGCT	53761						
Qy	358	gagcttggaaatctgaagaagctcatacgcttggatcttgtacaacaacatctaacccgg	417						
Db	53760	GACCCTGGAGAGCTCAAAACACTTATTAAGCTGGATCTGTATACAAACAATTAAGTGA	53701						
Qy	418	aaatcccatctcttccttggaaaattgaagtaactgttttttgcggctta	469						
Db	53700	ACCATACCGAAGAAGACTTGCGCAAGCTGAGTTCCCTTAGTATCTCTTGAGAGAA	53649						
	RESULT 14								
	LOCUS	ABO10698							
	DEFINITION	ABO10698	87434 bp	DNA	PLN	20-NOV-1999			
	ACCESSION	ABO10698							
	VERSION	ABO10698.1	GI:2828183						
	KEYWORDS	HTG.							
	SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_1lb:Mitsui P1 clone:MP112.							
	ORGANISM	Arabidopsis thaliana							
	AUTHORS	Eukariotae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; euclidoceles; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.							
	TITLE	Nakamura,Y							
	JOURNAL	Structural Analysis of Arabidopsis thaliana Chromosome 5. V							
	REFERENCE	Unpublished (1998)							
	AUTHORS	2 (bases 1 to 87434)							
	TITLE	Nakamura,Y							
	JOURNAL	Direct Submission							
	FEATURES	Submitted (28-JAN-1998) to the DDBJ/EMBL/GenBank databases.							
	source	Yasuhazu Nakamura, Kazusa DNA Research Institute, Laboratory of Gene Structure 2; 1532-3, Yama, Kisarazu, Chiba 292, Japan (E-mail:yinakamu@kazusa.or.jp, Tel:+81-438-52-3935, Fax:+81-438-52-3934)							
		Location/Qualifiers							
		1..87434							
		/organism="Arabidopsis thaliana"							
		/strain="Columbia"							
		/db_xref="taxon:3702"							
		/chromosome="5"							
		/clone="MP112"							
		/clone_1lb="Mitsui P1"							
	BASE COUNT	28656 a 15130 c 15592 g 28056 t							
	ORIGIN								
	Query Match	8.3%;	Score 73.8;	DB 7;	Length 87434;				
	Best Local Similarity	55.7%;	Pred. No. 2.4e-07;						
	Matches 141;	Conservative	0;	Mismatches	112;	Indels	0;	Gaps	0;
Qy	333	tctcaaaaaagagatccaaggaactatacctcttgagcttggaaatctgaagaagcttaa	382						
Db	52429	TGTATTCATAATGATCTTGAAGGTCACATCTCGAAGAAATGTTGATATGACGCTACTCT	52488						
Qy	383	tcaagttggatcgttacacacaaatctcacccgggaagaatcccatcttccttggaaaat	442						
Db	52489	CAGTTCCTGATCTTCCAACAAACAATTTCTCAGGTCGATTCCTGTTCTCCAAAGC	52548						
Qy	443	tgaagtaactggttttctgcgcttaacgaacaccgattgaacggtctctactagag	502						
Db	52549	TTGAAATGCCTTACTACTTGATAGCTTCAAGGAAACAATTCACGCGGTCTATCCCTGCCA	52608						



Qy	563	caattccagtaga	575
Db	52669	GCCTTAAGTGGCTTTCGCCTTCACACACTTCGATATCTCCGACAACTCTCTACAGGAA	52668
Qy	503	aacacaaagatattatgaagcttaagatgtgtgaatgcacaggaatgatttgtgtgaga	562

```
Search completed: June 24, 2000, 00:23:23
Job time: 46358 sec
```

RESULT 15

LOCUS	3176 bp	DNA	PAT	24-JUN-1998
DEFINITION	Arabidopsis thaliana cDNA encoding a protein involved in morphogenesis.			

REFERENCE 1 (bases 1 to 3176)  
AUTHORS Mitsuikawa,N. and Kobato,E.U.  
TITLE GENE CODING FOR PROTEIN CONTROLLING MORPHOGENESIS OF PLANT  
JOURNAL Patent: JP 1997056389-A 1 04-MAR-1997;

OS	COMMENT
Arabidopsis thaliana (thale cress)	
DN 1007055333-1/1	

BASE COUNT	927 a	654 c	659 g	936 t
ORIGIN				

Query Match	8.28;	Score 73.4;	DB 5;	Length 3176;
Best Local Similarity	54.78;	Pred. No. 3.9e-07;		
Matches 146;	Conservative	0;	Mismatches 121;	Indels 0;
			Gaps	0;

OY	309	gtcaatagtgtaactccacaaaacggagcttaagaacataaccccttcgagctggaaa	368
Db	1265	gacctaaccttaattgctgcacgaacacattatcaaaaggctcaatccgggttgagctatcgcg	1324
OY	369	tctgaagagcttaacgagtttgatgctgtacaaacaacatctcacccgggaanaatccacgc	428
Db	1325	tatcgctaaccttaatgactatgattgattctttccacacacaaagatataatgaaatgattctctc	1384
OY	422	ttcttgggaanaattgaatcactggttttttggcgcttaacgaaacagatataccgg	488
Db	1385	tttcctttggctgatttgagacatctcttcacaaagtgaacttcagtcgacaaacataatatacgc	1444
OY	488	tccctttccctgagaactcaacagttatcttaacagccttaagctgtgtgagtctcaaggaa	548
Db	1445	ttttaactcagaacgcgctttggaanaattcaagaagcatcattggaatagattcttcaaatpa	1504
OY	549	tgatttgctggaacaattccacagtaga	575
Db	1505	tgatatctcttgaccccaattccagagaga	1531



\_\_\_\_\_

7  
.  
.  
.

.

.

.

.

.

.

.

.

.

1



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:53:17 ; Search time 446.21 Seconds

(without alignments)  
501.270 Million cell updates/sec

Title: US-09-180-798-28

Perfect score: 894  
Sequence: 1 ggaacgattcaagcctccga.....gtcaacttgattcttattata 894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	894	100.0	894	1 V06589	Arabidopsis thalia
2	361	40.4	788	1 V06588	Arabidopsis thalia
3	361	40.4	1063	1 V06590	Arabidopsis thalia
4	356.2	39.8	981	1 V06587	Arabidopsis thalia
5	330	36.9	1106	1 V06586	Arabidopsis thalia
6	106	11.9	2089	1 V06591	Arabidopsis thalia
7	95.4	10.7	1814	1 V06571	Daucus carota SERK
8	73.4	8.2	3176	1 T62124	Arabidopsis thalia
9	71.6	8.0	4081	1 V06585	Arabidopsis thalia
10	66.4	7.4	3573	1 T06307	Partial tomato pat
11	66.4	7.4	6471	1 T06306	Tomato pathogen re
12	63	7.0	3541	1 V14522	CF-5 pathogen rest
13	63	7.0	6695	1 V06570	Daucus carota SERK
14	62.8	7.0	3979	1 V14518	CF-5 pathogen rest
15	62.8	7.0	3979	1 V14519	CF-5 pathogen rest
16	62.8	7.0	4123	1 V14523	CF-5 pathogen rest
17	62	6.9	1554	1 T31307	Tomato RRX gene cl
18	62	6.9	3293	1 X23532	Tomato Xa21 clone
19	61.2	6.8	3921	1 T31300	Rice Xa21 disease
20	61.2	6.8	13340	1 X23522	O. longistaminata
21	61.2	6.8	19639	1 X23524	O. longistaminata
22	59.6	6.7	7204	1 X23527	O. longistaminata
23	59.6	6.7	7204	1 X23527	O. longistaminata
24	54.8	6.1	8416	1 X23523	O. longistaminata
25	54.2	6.1	9424	1 X23525	O. sativa Xa21 gen
26	53.8	6.0	2880	1 O91450	Tomato Cf-9 CDNA
27	53.8	6.0	2880	1 T06304	Tomato pathogen re
28	53.8	6.0	3305	1 O91449	Tomato Cf-9 gene
29	53.8	6.0	3305	1 T06303	Tomato pathogen re
30	52.6	5.9	2192	1 X23530	Maize Xa21 gene DT
31	52.6	5.9	5740	1 X23526	O. longistaminata
32	52.4	5.9	5740	1 X00477	Arabidopsis thalia
33	52.2	5.8	3050	1 T06309	Tomato pathogen re
34	52.2	5.8	3089	1 T47877	Tomato pathogen re

35 48.8 5.5 3045 1 X23531 Maize Xa21 gene DT  
36 48.4 5.4 4104 1 X07356 Arabidopsis sterol  
37 45.2 5.1 3842 1 X23533 Tomato Xa21 clone  
38 40.6 4.5 2075 1 T49435 Tomato polygalactu  
39 38.8 4.3 9295 1 T62125 Arabidopsis thalia  
40 36.6 4.1 5361 1 T78868 P. falciparum live  
41 36.6 4.1 6152 1 T78867 P. falciparum live  
42 36 4.0 1117 1 Q42595 Sequence encoding  
43 36 4.0 1140 1 V26098 Mandarin DNA regio  
44 36 4.0 110000 1 V21209\_15 Continuation (16 o  
45 35.6 4.0 1380 1 V30822 Cytosolic glycerol

## ALIGNMENTS

```

RESULT 1
ID V06589 standard; cDNA to mRNA; 894 BP.
AC V06589:
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production: embryos;
KW plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT CDS 1..678
FT /tag= a
FT /note= "Shows high homology to SERK"

PN WO9743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR P-PSDB: W47021.
DR WPI: 98-086529/08.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 83-84; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences. 176 G; 285 T;
SQ Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;

Query Match 100.0%; Score 894; DB 1; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.4e-231;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggaacgattcaagcctccgaagggaagccttcaacgcgttgcgcgaagctatcaagat 60
DB 1 ggaacgattcaagcctccgaagggaagccttcaacgcgttgcgcgaagctatcaagat 60
QY 61 cgaagaatgtttcgaagtggtggaatccactctgttaacctgttactgtgttcat 120
DB 61 cgaagaatgtttcgaagtggtggaatccactctgttactgtgttactgtgttcat 120
QY 61 cgaagaatgtttcgaagtggtggaatccactctgttactgtgttactgtgttcat 120
DB 61 cgaagaatgtttcgaagtggtggaatccactctgttactgtgttactgtgttcat 120
QY 121 gtcacttgtaacacacacacacacacacacacacacacacacacacacacacac 180
DB 121 gtcacttgtaacacacacacacacacacacacacacacacacacacacacacac 180
QY 121 gtcacttgtaacacacacacacacacacacacacacacacacacacacacacac 180
DB 121 gtcacttgtaacacacacacacacacacacacacacacacacacacacacacac 180
QY 181 ggaacatctgtaacacacacacacacacacacacacacacacacacacacacacac 240
DB 181 ggaacatctgtaacacacacacacacacacacacacacacacacacacacacacac 240
QY 181 ggaacatctgtaacacacacacacacacacacacacacacacacacacacacacac 240
DB 181 ggaacatctgtaacacacacacacacacacacacacacacacacacacacacacac 240
QY 241 actctcttgccttctgattatcctgaacacattcatatcatcagtcacacataacatt 300
DB 241 actctcttgccttctgattatcctgaacacattcatatcatcagtcacacataacatt 300
QY 241 actctcttgccttctgattatcctgaacacattcatatcatcagtcacacataacatt 300
DB 241 actctcttgccttctgattatcctgaacacattcatatcatcagtcacacataacatt 300
QY 301 tgcattgagcatatagtggaactctacaaagagattcaaggaactatcactctgag 360
DB 301 tgcattgagcatatagtggaactctacaaagagattcaaggaactatcactctgag 360
QY 301 tgcattgagcatatagtggaactctacaaagagattcaaggaactatcactctgag 360
DB 301 tgcattgagcatatagtggaactctacaaagagattcaaggaactatcactctgag 360
QY 361 ctggaacatctgaagagctcaatcagtttgatctgtgaacaacaacatctcacaggaaa 420
DB 361 ctggaacatctgaagagctcaatcagtttgatctgtgaacaacaacatctcacaggaaa 420

```



```

|||||
Db 361 CTGGAATCTGAAAGCTATCATGTTGGATCTGTACAAACAATCTCACGGGAAA 420
Qy 421 atcccatctcttctgggaaatgaagtcacttgcttttggcgcttaagaaacga 480
Db 421 ATCCCATCTCTTCTTGGGAAATGGAAGTCACCTGTTTTTGGCGCTTAAGAAACCGA 480
Qy 481 ttgacccgctctctctcctgaagaacacagttatctcaagccttaagttctgtatgc 540
Db 481 TGAACCGGTCCTATCTCCTAGAGAACTACAGATTATTTCAAGCTTAAAGTTGTGAGTGC 540
Qy 541 tcaaggaaatgatctgtgtggaacaattccagtagaagaacctttgaacaattccatg 600
Db 541 TCAGGGAAGATTTGTGTGGAACAATCCAGAGAAAGACCTTTTGAACAACATTCCTATG 600
Qy 601 caaaacttggagaacaacctgagattggagggacagactactagctctgcagctat 660
Db 601 CAAACTTGGAGAACACTGAGATTGGAGGACCAAGACCTAGTGTGCTGCGACGCTAT 660
Qy 661 gacaccaattgacttaaaaaagaagltgaagaacctataaagaagaatgttagtgacct 720
Db 661 GACACCAATTGCACCTTAAAGAAAGATTGAGAACCTATTAAGAAAGATTAGGTGACCT 720
Qy 721 ttgaagaactctgtaccagttgtgttaactatataagacctgtttcagtata 780
Db 721 TGTAAAGACTCTGTACCAAGTGTGTAAATATATATAGAGCTGTGTTCATGTTATATA 780
Qy 781 tgaagcttggagagacagtaacttgcacatgtatgtatgttgaagaaagtgaagt 840
Db 781 TGAAGACTTGTGAGAGACAGTAATCTGCATGTATTGATTGTGTAAGAAAGTTGAAATG 840
Qy 841 agaatgtcttgaattgattgtgttctctatgtacttgaaattcttatta 894
Db 841 AGAATTGCTTGTAAATGGAATTGTGTCTTATGTAATGAAATTTCTATTA 894

RESULT 2
V06588 standard: cDNA to mRNA; 788 BP.
ID V06588;
AC 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT 2..664
FT CDS
PT /note= "shows high homology to SERK"
PN MO9743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: W47020.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 79-80; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 788 BP; 234 A; 191 C; 156 G; 207 T;

Query Match 40.4%; Score 361; DB 1; Length 788;
Best Local Similarity 72.8%; Pred. NO. 3.7e-88;
Matches 336; Conservative 0; Mismatches 110; Indels 91; Gaps 2;
Qy 15 ctccgaaggagacgctcttcacgcttcgctcgagactatcagatccagacaatgtgt 74
Db 91 CTCGGAAGAGAGATCGCTTTACGCTCTTCCGCGAGATTAAAGACGACCATGTCCT 150
Qy 75 tcaagattggatccaaactctgttaactctgttactgtgttcacatgtacactgtata 134

```

```

|||||
Db 151 CCAGAGCTTGAGTCCAAACTCTTGAATCCCTGTACCTGGTTCATGTCACTTAACA 210
Qy 135 acaccataagtcactcgtctgatttgggaattcaactatctgacatcagtaacc 194
Db 211 AGACAAACCGGTCACCTCGTGTGATTTGGGAATTCAAACCTCTGTGACATCTTGGGCC 270
Qy 195 tgaacttggaaagtgtgaacatttaacaatctgtatgaaatcactactcttgccttc 254
Db 271 TGAGTGTGGGAAGCTGAACATTTCAGTACT----- 304
Qy 255 tgattctcgaaaaaacttaactatcagtcacacataataacatttgcgtgtaata 314
Db 304 ----- 304
Qy 315 tagtgactctcaaaaaacagatctcaaggaactaactctcgtgacttgaacatctgaa 374
Db 304 ---AGAGCTCTCAAAAAACATCAACAGAGACTATACCTTCCGAATCTGAAATCTGAA 360
Qy 375 gagcttaactcagttgtgactctgacacacaaatctcacgggaaatcccatctctt 434
Db 361 GAATCTCATCAGCTTGATCTGTACAAACAATCTTACAGGATATGTTCCACTCTTT 420
Qy 435 gggaaaaattgaagtcactgttttttgggcttaacgaaaaacgacttgcagctat 494
Db 421 GGGAAATTCAGTCTCTGTCTTTTACGCTTATATGCAACGATTTGACGGGCCAAT 480
Qy 495 tcttagagaactcagattatctcaagccttaagttgtatgtctcgaagaatgatct 554
Db 481 CCTTAGAGCACTCAGTCGCAATCCAGCCTTAAAGTTGTGATGTCTAAGCAAGATTT 540
Qy 555 gtgtgaacaactcagtagaaggaacctttgaacacattcctatgcaaaacttggaa 614
Db 541 GTGTGGAAACATCCAAACAAGGACCTTTGCTCAATCTCTTACGAACATTGAGAA 600
Qy 615 caacttgaatttggaggaacagactactaggtcttgcagactatgacccaattggac 674
Db 601 CAACCCGAGGTGGAGGAGACCGGAATTACTCGTCTTGCACGTAAGACATAATGAC 660
Qy 675 ttaaaagaagttggaagacacttaagaagaa-tgttgggtgactctgttaagactgt 733
Db 661 CTGAAAAAATTTGGCAAACTGAAAAATGAAGATTGGGGGGTGCCTGTATAAGAACTT 720
Qy 734 taccagttgttgaact 752
Db 721 CACCACTTATCAAAATATC 739

RESULT 3
V06590 standard: cDNA to mRNA; 1063 BP.
ID V06590;
AC 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT 106..762
FT CDS
PT /note= "shows high homology to SERK"
PN MO9743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR WPI: 98-086529/08.
DR P-PSDB: W47022.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 86-88; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.

```



Seq	Sequence	1063 BP;	313 A;	242 C;	206 G;	302 T;
Query Match	40.4%;	Score 361;	DB 1;	Length 1063;		
Best Local Similarity	72.8%;	Pred. No. 4.1e-88;				
Matches 538;	Conservative	0;	Mismatches 110;	Indels 91;	Gaps	2
QY	15	ctccgaaggagacgcctctccacgcgcttcgcgaggtcctacagatccagacaattggtg	74			
Db	189	CTCGGAAGGAGATGCTCTTACGCTCTTGGCGGAGTTTAACAGATCCGGACATGTTCT	248			
QY	75	tccagattggatcccaactctgttaacccctgtactctgttcaatgtaactgttaaca	134			
Db	249	CCAGAGCTGGGATCCACACTCTGTATAATCTGTACCTGGTTCATGCTACACCTGTATCA	308			
QY	135	acacacatcaagtacacgcctcgtgatttgggagattcaactatctcgtgacatctag	194			
Db	309	AGACAACCGCGCTACACGCTGTGGATTGGGGAAATTCAAACTCTGTGACACTTGGCC	368			
QY	195	tgaacttgggaagcttggacaattcaatatctgtatgaaatcaatcaactctttgcctt	254			
Db	369	TGACCTTGGGAACCTTGAACTTTAAGTATCT-----	402			
QY	255	tgtatctctaaacattacattacagtcacacataaacaatttgcctttagtgcata	314			
Db	402	-----	402			
QY	315	tactgaacttaaaaaagagatttaaggaactataactctcgtgacttgaactctgaa	374			
Db	402	---AGAGCTCTACAAAAACAATCCAAAGAACTTATACCTTCCGAATCTGGAAATCGAA	458			
QY	375	gagctaatcagtttgatctgtacacacacaactctcacgsgaanaatccatctctt	434			
Db	459	GAATCTCATCAGCTGTGATCTGTACAAACAATCTTAAGGAGATGTCCACTCTTT	518			
QY	435	ggsgaaattaaagtcactgttttttggggttaacgaaaaacgatttgaccgtctcat	494			
Db	519	GGGAAATTAAGTCTCTGTCTTTTAAAGGCTTATGACAACCATTTGACGGGGCAAT	578			
QY	495	tcttagaagaaactacagtttaatttcaagcctttaaagtgttgatgctcagsgaa	554			
Db	579	CCCTAGAGCACTCAACGCAATCCCAAGCCTTAAAGTGTGATGCTCAACCAATGATTT	638			
QY	555	gtctggaacaattccagttgaaggaacttttgaacaacatctcctatgsgaaacttsgaa	614			
Db	639	GTGTGGAACATCCACACACGAGACCTTTGTCTCACTCTTTTACGAACCTTTGAGA	698			
QY	615	caacttgaatttggagggaccagaactactagtcttgcgactatgacacaaattgcac	674			
Db	699	CAACCCAGAGTTGGAGGAGCCGGAATTAATCTCGTCTTCAACCTTCGACATTAAC	758			
QY	675	ttaaaagaagtttgaagaacctataagaagaa-gtttagtgggaactgttaagaacctg	733			
Db	759	CTGAAAAAATGGCAAAACCTGAAAAATGAAGATTGGGGGTGACCTGTGAAGAACCTT	818			
QY	734	taccagtgcttgaataac	752			
Db	819	CACCACCTTATCAAAATATC	837			
RESULT	4					
V06587						
ID	V06587	standard; cDNA to mRNA; 981 BP.				
AC	V06587;					
DE	03-AUG-1998	(first entry)				
DE	Arabidopsis thaliana SERK LRR homologous EST clone.					
KW	receptor kinase; apomixis; apomixis; seeds; production; embryos;					
KW	plant breeding; lenceine-rich repeat; ss.					
OS	Arabidopsis thaliana.					
PH	Key	location/Qualifiers				
FT	CDS	104..760				
FT		/*tag= a				
FT		/note= "shows high homology to SERK"				

PD MN9743427-11.  
 PD 20-NOV-1997.  
 PR 13-MAY-1997; E02443.  
 PR 14-MAY-1996; GB-010044.  
 PA (NOVS) NOVARTIS AG.  
 PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
 PI WPI; 98-086529/08.  
 DR P-PSDB; M47019.  
 PT Production of apomictic seeds - useful in plant breeding  
 PS Claim 28; Pages 75-77; 123pp; English.  
 CC The sequence is that of an EST clone showing high homology to  
 CC SERK LRR (leucine-rich repeat) sequences.  
 SQ Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;

Query Match 39.8%; Score 356.2; DB 1; Length 981;  
 Best Local Similarity 72.4%; Pred. No. 7.8e-87;  
 Matches 535; Conservative 0; Mismatches 113; Indels 91; Gaps 2;

QY 15 ctccgaaggagacgctctctcaecgcctccgcggagcttaccagatccagaatgtgt 74  
 DB 187 CTCGGAAGGAGATGCTCTTACGGCTCTTGGCGGAGTTGACAAATCCAAACATCTCT 246  
 QY 75 tcaagatgtagatcccaactctgttaacacctgtgactggttcatatgcactgtaata 134  
 DB 247 CCAAGAGCTGGGATCCAACTCTGTAAATCTGTAACTGGTGTCAATGTCAACGTAAACA 306  
 QY 135 acaccatcaagtctactcgtctgtgattggtgggaattcaaatctatcggacattagacc 194  
 DB 307 AGAATAACCGGCTACCTGCTGTGGAATTTGGGAAATTCAAACCTCTGGAATCTTCCGCC 366  
 QY 195 tgaactgtagaagcttgaaacttacaatcatcgtatgagaatcacactcttgcctt 254  
 DB 367 TGACCTTGGGAAGCTTGAAACATTTACAGATCT----- 400  
 QY 255 tgaattctgaaacattacattcatcagttacacataataacatttgcctttagtcata 314  
 DB 400 ----- 400  
 QY 315 tagtgaacttacaataacgagattcaagaaactataactctgacttggaaattgaa 374  
 DB 400 ---AGACCTCTCAAAAACAAACATCCAAAGAACTAATACCTTCGGAATCTGAA 456  
 QY 375 gagctctatagcttggatctgtacaacaacaattctcacccggggaataatccactctctt 434  
 DB 457 GAACTCTATATAGCTTGATGTGATCTGTAAACAACAATCTTACAGGGAATTAATCCCACTCTT 516  
 QY 435 gggaataattgaatcacttgttttttgcgggttaacgaaacggatctgacggtctcat 494  
 DB 517 GGGAAATTTGAATGTCCTGTCTTTTACGGCTTTAATACAAACGATTTGACCGGTCCAT 576  
 QY 495 tctcagaagaactcacagttatttcaagccttaagcttgaagtgttgaatctcagaagaatatt 554  
 DB 577 CCGTAGAGCACTACGCGCAATCCCAAGCCTTAAATTTGTTGACCTTCACGACATATATT 636  
 QY 555 gtgtgagacattccagtagaaggaaccttltgaaacattcctatgtgaaactgtaga 614  
 DB 637 GTGTGGAACATCCCAACAAGGACCCCTTTCTCATCTTCTTTAAGAAACTTTGAGA 696  
 QY 615 caacctgagttgagaggaacagaactactagctcttcgagactatgaaaccaaatgtgac 674  
 DB 697 CAACCCGAGATTTAGAGGAGCCGGAAATTAATCTGCTTTCAGAGTACGACTAATGAC 756  
 QY 675 ttaaaagaagaattgaaagaaactataaagaaga-tgttaagtgacactgtgaagaactctg 733  
 DB 757 CTGAAGAACATGGCAAAACCTGAAATTAAGAAATTTGGGGGTGACCTGTAAAGAACCTT 816  
 QY 734 taccgaatgttggtaaatc 752  
 DB 817 CACCACTTTATCAAAATATC 835



V06586  
ID V06586 standard; cDNA to mRNA; 1106 BP.  
AC V06586;  
DT 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK LRR homologous EST clone.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KM plant breeding; leucine-rich repeat; ss.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 142..798  
FT /tag= a  
FT /note= "shows high homology to SERK"  
PN MO9743427-A1.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS) NOVARTIS AG.  
PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
P-PSDB: W47018.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 28; Pages 71-73; 123pp; English.  
CC The sequence is that of an EST clone showing high homology to  
CC SERK LRR (leucine-rich repeat) sequences.  
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;

Query Match 36.9%; Score 330; DB 1; Length 1106;  
Best Local Similarity 71.9%; Pred. No. 9.2e-80;  
Matches 532; Conservative 0; Mismatches 115; Indels 93; Gaps 4;  
QY 15 ccccgaaaggagcgcctcttcacgcgctccgcggagcttaccagtcacgaatgtgt 74  
DB 225 cttccgaagagatgctcttcttaccgcttccggagttgacagatccagacatgctt 284  
QY 75 tcaagtttggaatccaactctgttaactcctgttactgtgttcttcaatgttaac 134  
DB 285 ccagagctggagatccaaactctgttaactcctgttactgtgttcttcaatgttaac 344  
QY 135 aaacatcaagctcgtctgtggttggttggttggttggttggttggttggttggt 194  
DB 345 agacaaacccgctacgtctgtggttggttggttggttggttggttggttggttggt 404  
QY 195 tgaacttggaagcttgaaacattacaatctgtatgtgaatcattcatttgcctt 254  
DB 405 tgaagcttggaagcttgaaacattacaatctgtatgtgaatcattcatttgcctt 438  
QY 255 tgattatctgaaacatttaccatcagtcacacataacatttgccttgagcata 314  
DB 438 ----- 438  
QY 315 tagtgaactctaaacagagatccaaggaactatacttctgagcttggaatctgaa 374  
DB 438 ----AGAGCTCTCAAAACAAACATCCAGGAATCTACTCTCGAATCTGGAAATCTGAA 494  
QY 375 gagctcaatcagtttgatctgtacaaacaaacatcaccgggaaatccatcttctt 434  
DB 495 GAATCTCATCAGCTTGATCTGTACAAACATCTTACAGAGGAGTATCCCATCTTCT 554  
QY 435 gggaaatgaagctcgtgttttttggttggttggttggttggttggttggttggttg 494  
DB 555 GGAATAATGAAGTCTCTGCTTTTACGGCTTAATGACAAACCAATGACCGGTCC-AA 613  
QY 495 tccatagaagactcaagttatttcaagcc-ttaaagtttgatgttccaaggaatgatt 553  
DB 614 TCTTAAGAGACTCAGCGCAATCCCAAGCCTTAAAGTTGTGAGCTTCACATGATATT 673  
QY 554 tgtgtgaaacattccagtagaagacatttgaacacattccattgcaaaacttgaga 613  
DB 674 TGTGTGGAGCAATCCCAAAACGAGACCTTGTCTCATCTTCTTAAAGAACTTTGAGA 733  
QY 614 acaaaccttgagtgaggagacgaactactagctcttcgagctcatgacacaaattca 673  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 734 ACAACCCGAGATGGAGGACCGGAAATTAATCTGCTTGCAGCTACGACACTACTGCA 793  
QY 674 ccttaaaagaagttgaagaacctataagaagaa-tgttagtgaccttgtaagactct 732  
DB 794 CCTGAACAACTGCGCAAAACCTGAATAATGAGAAATTTGGGGGTGACCTTGAAGAACACT 853  
QY 733 gtaccacagtggttgtaactc 752  
DB 854 TCACCACTTATCAAAATATC 873  
RESULT 6  
V06591  
ID V06591 standard; cDNA to mRNA; 2089 BP.  
AC V06591;  
DT 03-AUG-1998 (first entry)  
DE Arabidopsis thaliana SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KM plant breeding; ds.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 195..2072  
FT /tag= a  
FT /product= SERK protein  
PN MO9743427-A1.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GB-010044.  
PA (NOVS) NOVARTIS AG.  
PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
P-PSDB: W47023.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 27; Pages 91-95; 123pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilize a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;  
Query Match 11.9%; Score 106; DB 1; Length 2089;  
Best Local Similarity 53.5%; Pred. No. 2.2e-19;  
Matches 257; Conservative 0; Mismatches 205; Indels 18; Gaps 1;  
QY 19 gaagggaagctcttaccgcgcttgccggagcttatcagatccagaaatgttttag 78  
DB 279 GAAGGAGATGCTTTCATCTTGAAGGTTACTCTAGTTGACCAAAATGCTTTGAG 338  
QY 79 agttggatcaactcttgtaactcttgtaactgtttcattgtaactgtatacaac 138  
DB 339 AGCTGGATCTCTACGATGGAATCTTGCACATGTTGATCTGCAACAACAGAC 398  
QY 139 catcaagctacgtctgtgatttgggaaattcaactatctggaacatctagctgaa 198  
DB 399 AACAGTGTCAATAGAGTGAATTTGGGGAATGCAAGTATCTGCGCATTAATGATTCAGAG 458  
QY 199 ctgggaagcttgaacattacaatatctgtatgaaatcattacccttcttgctttag 258  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||



Db	459	CTTGGTCTCTCAAAATTTCAGATTATGGA-----GCTTTACAGT	5001
Qy	259	tatctgaanaacattacatttagtcacacataaacatttgcctttgagtaatacagt	318
Db	501	AACAACATACACGCGCCCAATTCATGTAATCTGGAAATCTGACAACTATGATGATTG	5605
Qy	319	gaacctacaaaaacagagattcaagaactatacctctcagcttggaaatctgaagagt	378
Db	561	GATCTTAACTTAAACAGCTTCCGCGCTCATTTCCGGAATCATTTGGGAAAGCTTTCAAG	6205
Qy	379	ctaatcagtttgagctctgcaacaacaacatctccagggaagaatcccatctctttggga	438
Db	621	CTGAGATTTCTCCGGCTTAAACAACAACAGCTCATCGGGTCAATTCATATGCACTGACC	6805
Qy	439	aaattgaagtcaactggttttttttgcgcttaagaagaaaccgattgacggctccatctcc	498
Db	681	AATATTAATCTACCTTCAAGCTGTTAGATATATCAATATACAGACTCTCTGGTCACTTCT	7405

RESULT	7
AC	V06571 standard; cDNA; 1814 BP.
AD	V06571.
AE	03-AUG-1998 (first entry)
DE	Daucus carota SERK gene.
DE	receptor kinase; apomixis; apomictic; seeds; production; embryos;
KW	plant breeding; ss.
OS	Daucus carota.
FT	key
FT	location/qualifiers
FT	94..1755
FT	/*tag- a
FT	/product- SERK protein
PN	MO9743427-A1.
PD	20-NOV-1997.
PF	13-MAY-1997; E02443.
PR	14-MAY-1996; GB-010044.
PA	(NOVS ) NOVARTIS AG.
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR	WPI: 98-086529/08.
DR	P-PsDB: M47013.
PT	Production of apomictic seeds - useful in plant breeding
CC	Claim 21: Pages 47-51; 123pp; English.
CC	The sequence is that encoding SERK, a putative receptor kinase.
CC	It may be used as part of a method of producing apomictic seeds
CC	comprising: (a) transforming plant material with a nucleotide
CC	sequence encoding a protein which in active form in a cell or
CC	cell membrane renders the cell embryogenic; (b) regenerating
CC	the transformed material into plants or carpel-containing
CC	plant parts; and (c) expressing the sequence in the vicinity
CC	of the embryo sac. The apomictic seeds and embryos thus produced
CC	can be developed into plant progeny. This is useful in plant
CC	breeding programs. Controllable and reproducible apomixis provides
CC	many advantages in plant improvement and cultivar development in
CC	the case that sexual plants are available as crosses with the
CC	apomictic plant. Apomixis provides for true-breeding, seed
CC	propagated hybrids and could shorten and simplify the breeding
CC	process so that selfing and progeny testing to produce and/or
CC	stabilise a desirable gene combination could be eliminated.
CC	Apomixis allows plant breeders to develop cultivars with
CC	specific stable traits for such characteristics as height,
CC	seed and forage quality and maturity.
CC	Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

Query Match	10.7%	Score 95.4	DB 1	Length 1814
Best Local Similarity	56.1%	Pred. No. 1.5e-16		
Matches 180; Conservative	0	Mismatches 141	Indels 0	Gaps 0

Oy 319 gaactctacaaaacgagattcaaggaactatcccttcgagcttgyaaactgaaagt 378  
 |||||  
 Db 178 GAGCTTACGCAATACATAGTGAGACCAATCTCTAGTCATCTTGGGAACCTGACAAT 237

QY	379	ctatcaagtttgagctcttaacaaacaatctacccgggaataatcccatctcttttggga	438
Db	238	TTGGTAGCTGTGGACCTATACATGAATAAGCTTCCTCGACCTATACCGGACATTATGA	297
QY	439	aaattgaagctactggttttttttgcggttaacgaataacggatgacggtctctctct	498
Db	298	AACTTACAAGGCTAAGATCTTGTGGTGTCTACACAACAGGCTCTCTGTGCTCAATTCA	357
QY	499	aggagactcaagctattcttaacgacttaagaattgtgagtcctcagggaatatttgt	558
Db	358	ATGTCACTGACATATATTCACAACCTTCAACCTCTCGATTTATCAACAACATGGCTATCA	411
QY	559	ggacaacatccagtagaagagaccttttgaacacatctctctgtcaaaactttgagaacac	611
Db	418	GGACCGGTACCGGATATATGGCTATTTTCTTTTACACCTATACAGTTTTCGCAATAT	477
QY	619	ctagagatggagggaccagaa	639
Db	478	TTGATATTATGTGAGCCCGTA	498

RESULT	8
ID	T62124
AC	T62124 standard; cDNA to mRNA; 3176 BP.
DT	T62124;
DE	10-JUN-1997 (first entry)
DE	Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.
KW	Plant; morphogenesis; regulation; short; stem; alteration;
KW	inflorescence; extraneous; gene; expression; transformation;
KW	increase; control; form; length; ds.
OS	Arabidopsis thaliana.
FS	Key
FS	Location/Qualifiers
FT	cds
FT	51..2961
FT	/*tag= a
FT	/note= "plant morphogenesis regulatory protein"
FN	
PD	J09056382-A.
PD	04-MAR-1997.
PF	24-AUG-1995; 216187.
PR	24-AUG-1995; JF-216187.
PA	(MITSU-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA	(CHIK-) ZH CHIKUY KANKYO SANGYO GIJINSU KENKYU.
DR	WPI: 97-206629/19.
DR	P-RSDS: M13408.
PT	DNA encoding plant morphogenesis regulatory protein - useful to
PT	yield plants with short stems or altered inflorescence
PS	Claim 1; Pages 6-10; 17pp; Japanese.
CC	The present sequence encodes an Arabidopsis thaliana plant
CC	morphogenesis regulatory protein (MRP), which can be used to yield
CC	a plant with, e.g. short stems or altered inflorescence. The MRP
CC	acts on a plant at a specific site for a specific period, and can
CC	therefore be used to regulate extraneous gene expression in a
CC	plant. The MRP's cDNA or genomic DNA can be used to transform a
CC	plant to increase its MRP expression, and therefore control the
CC	form (particularly stem length) of the plant.
SQ	Sequence 3176 BP: 927 A; 654 C; 936 T;

Query Match	8.28;	Score 73.4;	DB 1;	Length 3176;
Best Local Similarity	54.78;	Pred. No. 1.6e-10;		
Matches 146;	Conservative	0;	Mismatches 121;	Indels 0;
			Gaps	0

QY 309 gtcatatagtagaacctctacaaaaacgagattccaaggaaactataaccttctgaagcgttgaaa 368  
| | | | | | | | | | | | | | | |  
Db 1265 GACTTACCTTAATGTGTCAGCAACAATATCAAAGTGCCAAATCCCGGTGAGCATCTCG 1324

Qy 369 tcgaagagtctaatacagtttgtagctgtacaaacaacatctcacgcggtaaaaattccatc 428  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1325 TATGGTACTTAGATACATTGAGTCTTCACAAACAAGATAAATGGAATCATTTCTTC 1384

```
Oy   429 ttccttggaagaattgaatcaactgttttlttcgcgcttaacgaanaaccgatggcgg    488
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1385 TTCCTTGTCATTTGGAGCATCTTCCAGAAGATGACCTGTAGTAGAATAATCATTAACTCG 1444
```



Oy		489	tccattcttcagagaactcaacagtattcttaaggcctaaggttgtagtcccgagaa	548
Db	1445	TGTAAGTTCACAGGCACCTTTGGAAATCTAAGAAGCATCATGGAATAGATTCTTCAATAA	150	
Oy	549	tgatttgtytggaacaattcccagtaga	575	
Db	1505	TGATATCTCTGTGCCCAATCCAGAGA	1531	
RESULT	9			
V06585				
AC	V06585	standard; DNA; 4081 BP.		
DT	03-AUG-1998	(first entry)		
DE	Arabidopsis thaliana SERK gene.			
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;			
OS	plant breeding; ds.			
OS	Arabidopsis thaliana.			
FT	Key	location/qualifiers		
FT	CDS	3696..6620		
FT		/tag= a		
FT	Intron	/note= "contains introns"		
FT		3731..3802		
FT		/tag= b		
FT	Intron	/number= 1		
FT		3851..3979		
FT		/tag= c		
FT	Intron	/number= 2		
FT		4124..4211		
FT		/tag= d		
FT	Intron	/number= 3		
FT		4284..4357		
FT		/tag= e		
FT	Intron	/number= 4		
FT		4430..4528		
FT		/tag= f		
FT	Intron	/number= 5		
FT		4642..4757		
FT		/tag= g		
FT	Intron	/number= 6		
FT		4890..4967		
FT		/tag= h		
FT	Intron	/number= 7		
FT		5295..5803		
FT		/tag= i		
FT	Intron	/number= 8		
FT		6197..6339		
FT		/tag= j		
FT		/number= 9		
PN	WO9743427-A1.			
PD	20-NOV-1997.			
PF	13-MAY-1997.	E02443.		
PI	14-MAY-1996; GB-010044.			
PA	(NOVS ) NOVARTIS AG.			
PI	De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;			
DR	WPI: 98-086529/08.			
DR	P-PsDB; W47017.			
PT	Production of apomictic seeds - useful in plant breeding			
PS	Claim 26; Pages 64-67; 12pp; English.			
CC	The sequence is that encoding SERK, a putative receptor kinase.			
CC	It may be used as part of a method of producing apomictic seeds,			
CC	comprising: (a) transforming plant material with a nucleotide			
CC	sequence encoding a protein which in active form in a cell or			
CC	cell membrane renders the cell embryogenic; (b) regenerating			
CC	the transformed material into plants or carpel-containing			
CC	plant parts; and (c) expressing the sequence in the vicinity			
CC	of the embryo sac. The apomictic seeds and embryos thus produced			
CC	can be developed into plant progeny. This is useful in plant			
CC	breeding programs. Controllable and reproducible apomixis provides			
CC	many advantages in plant improvement and cultivar development in			
CC	the case that sexual plants are available as crosses with the			
CC	apomictic plant. Apomixis provides for true-breeding, seed			
CC	propagated hybrids and could shorten and simplify the breeding			

CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and storage quality and maturity.  
SQ Sequence 4081 bp; 1120 A; 770 C; 785 G; 1406 T;

Query Match	8.0%;	Score 71.6;	DB 1;	Length 4081;
Best Local Similarity	54.4%;	Pred. No. 5.2e-10;		
Matches 186;	Conservative	0;	Mismatches 129;	Indels 27;
			Gaps	1

Qy	217	ttacaatactctgfatg-----	-----	-atacaactctttg	249
Db	2075	TTTGACGATTTTGTAAAGTCCATTATGACATGATGCTTTAAACAAAACAAATCCAGATTTG			2133
Qy	250	ccctttgattatctgaaacaattacattacagtcacacatacaatattgctttgag			309
Db	2135	ACGAAAGAAGCACTGGAGTTACCTTTTGTGAATGGAATCTTTTAAACAGTTTCTTATT			2194
Qy	310	tcataatgtyaaactctacaaaacgagatccaaggaactatacccttctgagcttggaaat			369
Db	2195	TCCTTACAGGAGCTTTACATATCAATCAACATPAACTGGCCGCAATTCCTGATGATTTGTGAAAT			2255
Qy	370	ctgaagagtcacatcaagttctgatacaacaacaactctcaaccggygaaatcccatct			429
Db	2255	CTGACAAACTTAGTAGATTTTGGATCTTTACTTAAACAGCTTCCGGGTCCATTCGCGAA			2314
Qy	430	tctttggyaaaaattgaaagtcaacttgtttttcttgcgcttaac			471
Db	2315	TCATTTGGAAAAGCTTTCAAAGCTGAGATTTCTGTGAGTATAC			2356

RESULT 10

ID T06307 standard; cDNA: 3573 bp.

AC T06307;

DT 14-APR-1996 (first entry)

KW Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.

DE Pathogen resistant; Cf-2.2; tomato; C.fulvum; Avr 4; Avr 9; fungal;

RV leaf mould; variegation; ss.

OS Lycopersicon esculentum.

OS W05531564.X2.

PN 23-NOV-1995.

PD 23-NOV-1995.

PF 11-MAY-1995; G01075.

PR 11-MAY-1994; GB-009394.

PR 23-DEC-1994; WO-G02812.

PR 31-MAR-1995; GB-006658.

PR 07-APR-1995; GB-007232.

PA (GATS-) GATSBY CHARITABLE FOUND.

PI Hammond-Kosack KE, Jones DA, Jones JDC;

DR WPI: 96-010949/01.

P-PSDB; R85299.

PT Increasing plant pathogen resistance by induction of variegation -

PT may lead to acquired resistance to a broad range of pathogens.

PS Claim 9; Page 85-87; 13pp; English.

CC T06307 is a tomato pathogen resistance gene Cf-2.2 partial cDNA clone

CC in a new method this gene is expressed highly in genetic constructs

CC which may be used to impart a broad range of pathogen resistance, by

CC induction of variegation, to transgenic plants (or parts or propagules

CC of plants) containing such constructs. Cf-2.2 imparts resistance to the

CC disease caused by the leaf mould fungal pathogen Cladosporium fulvum.

CC C.fulvum contains avirulence (Avr) genes that confer recognition by

CC plants containing Cf genes, leading to the activation of host defence

CC mechanisms to attack the disease. 654 C; 664 G; 1223 T;

SO Sequence 3573 bp; 1032 A; 654 C; 664 G; 1223 T;

7.48; Score 66.4; DB 1; Length 3573;



Best Local Similarity 54.0%; Pred. No. 1.2e-08;  
Matches 136; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
OY 320 aactcacaacagagattcaaggaactataccttcgtgagcttggaaatcgaagatc 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1230 ATCTCTACAAATATACAGCTTTCGCTCTATTCTGAAGAAATAGTACTTGAGTCTC 1289
OY 380 taatcagtttgatctgtacaacaacatctcaccgggaaatcccatctctctggaa 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1290 TTAATTATCTAGATTGAGTAATACATCATTAATGATTTATTCCTGCTCATTTGGCA 1349
OY 440 aattgaagtcactgttttttcgagcttaacgaacacgattgacggctcattccta 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1350 ATATGACCAACTGGCTTTTGTCTTTTATGAAATACACTTGCTACTCTGTCTCTG 1409
OY 500 gagaactcacagttattcaagccttaagttgtatgtctcagggaatgttgtgtg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1410 AAGAAATAGGTACCTAAGTCTCTTAATGTCCTGATTGAGTGAGAAATGCTTTAATG 1469
OY 560 gaacaattccag 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1470 GCTCTATTCTCTG 1481
```

## RESULT 11

```
T06306
ID T06306 standard; DNA; 6471 BP.
AC T06306;
DT 14-APR-1996 (first entry)
DE Tomato pathogen resistance gene Cf-2.1.
KW Pathogen resistant; Cf-2.1; tomato; C.fulvum; Avr 4; Avr 9; fungal;
KM leaf mould; variegation; ds.
OS Lycopersicon esculentum.
FH Key location/Qualifiers
FT 5'utr 1..1676
FT cds /*tag- a
    1677..5014
    /*tag- b
    /product- immature_Cf-2.1_protein
FT signal_peptide /*tag- c
    1677..1745
FT mat_peptide /*tag- d
    1746..5011
FT 3'utr /*tag- e
    5015..6471
FT W09531564-A2.
PN 23-NOV-1995.
PR 11-MAY-1995; G01075.
PR 11-MAY-1994; GB-008394.
PR 23-DEC-1994; WO-G02812.
PR 31-MAR-1995; GB-006658.
PR 07-APR-1995; GB-007232.
PA (GATS-) GATSBY CHARITABLE FOUND.
PI Hammond-Kosack KE, Jones DA, Jones JDG;
    WPI: 96-010949/O1.
DR P-PSDB; R85298.
FT Increasing plant pathogen resistance by induction of variegation -
    may lead to acquired resistance to a broad range of pathogens.
PS Claim 9; Page 80-83; 131pp; English.
CC T06306 is the tomato pathogen resistance gene Cf-2.1. In a new method
    this gene is expressed highly in genetic constructs which may be used
    to impart a broad range of pathogen resistance, by induction of
    variegation, to transgenic plants (or parts or propagules of plants)
    containing such constructs. Cf-2.1 imparts resistance to the disease
    caused by the leaf mould fungal pathogen Cladosporium fulvum.
CC C.fulvum contains avirulence (Avr) genes that confer recognition by
    CC plants containing Cf-genes, leading to the activation of host
    CC defence mechanisms to attack the disease.
SQ Sequence 6471 BP; 2073 A; 1106 C; 1122 G; 2170 T;
```

Query Match 7.4%; Score 66.4; DB 1; Length 6471;  
Best Local Similarity 54.0%; Pred. No. 1.6e-08;

Matches 136; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
OY 320 aactcacaacagagattcaaggaactataccttcgtgagcttggaaatcgaagatc 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2914 ATCTCTACAAATATACAGCTTTCGCTCTATTCTGAAGAAATAGTACTTGAGTCTC 2973
OY 380 taatcagtttgatctgtacaacaacatctcaccgggaaatcccatctctctggaa 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2974 TTAATTATCTAGATTGAGTAATACATCATTAATGATTTATTCCTGCTCATTTGGCA 3033
OY 440 aattgaagtcactgttttttcgagcttaacgaacacgattgacggctcattccta 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3034 ATATGACCAACTGGCTTTTGTCTTTTATGAAATACACTTGCTACTCTGTCTCTG 3093
OY 500 gagaactcacagttattcaagccttaagttgtatgtctcagggaatgttgtgtg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3094 AAGAAATAGGTACCTAAGTCTCTTAATGTCCTGATTGAGTGAGAAATGCTTTAATG 3153
OY 560 gaacaattccag 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3154 GCTCTATTCTCTG 3165
```

## RESULT 12

```
V14522
ID V14522 standard; DNA; 3541 BP.
AC V14522;
DT 20-MAY-1998 (first entry)
DE Cf-5 pathogen resistance gene clone Hcr2-5B gene.
KW Tomato; Cf-5 pathogen resistance gene; Cladosporium fulvum-5;
    tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon pimpinellifolium.
FH Key location/Qualifiers
FT CDS 603..3002
FT /*tag- a
    /product- Cf-5 pathogen resistance gene clone Hcr2-5B
FT PN W09743429-A1.
FT 20-NOV-1997.
PR 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1996; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
    WPI: 98-008895/O1.
DR P-PSDB; W41311.
FT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
    for production of transgenic plants resistant to pathogens e.g.
PT tomato leaf mould C.fulvum in tomatoes
PS Disclosure; Fig 5; 75pp; English.
CC This sequence is a clone of the polynucleotide of the invention, which
    is able to confer pathogen resistance on a plant. The polynucleotide of
    CC the invention is one of two tomato Cf-5 gene variants, which offer
    CC resistance against the pathogen Cladosporium fulvum-5. Transgenic plants
    CC can be produced by incorporating the gene into plant cells and
    CC regenerating plants from the cells; asexually or sexually produced
    CC offspring can also be subsequently produced. Expression of the gene in
    CC plant cells can confer pathogen resistance on a plant e.g. to tomato leaf
    CC mould (C. fulvum) in tomatoes. Oligonucleotides with sequences
    CC complementary to the gene or fragments of it, are useful in anti-sense
    CC techniques to reduce gene expression. The nucleic acids/polynucleotides
    CC are useful as hybridisation probes to identify other genes/fragments
    CC conferring pathogen resistance on plants e.g. Phytophthora resistance in
    CC potatoes. Homologues between Cf-5 and Cf-9 may be used to identify
    CC further resistance genes of this class.
SQ Sequence 3541 BP; 1123 A; 630 C; 612 G; 1176 T;
```

Query Match 7.0%; Score 63; DB 1; Length 3541;  
Best Local Similarity 52.5%; Pred. No. 1e-07;  
Matches 138; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

OY 309 gtcatatagtgaaactcacaacagattcaaggaactataccttcgtgagcttggaa 368  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||







Query Match 7.0%; Score 62.8; DB 1; Length 3979;  
 Best Local Similarity 53.2%; Pred. No. 1.2e-07;  
 Matches 133; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

QY 320 aactctacaaaagagattcaaggaaactatatacctcttgagcttggaatctgaagagtc 379
    ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2107 ATCTTTACAAATAATACAGCTTTCGCTCTATTCCTGCTTATTCGCAATATAGAAATTC 2166
QY 380 taatcaagttggatctgttacacaacatctcacccgggaaatccatctctcttg9aa 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2167 TGCAAACTCTGTTTCACGATTAACGATCTCATTTGGGAAATTCCTTCATTGTGTGCA 2226
QY 440 aattgaagtcacttgcttcttgcggtctaacgaaacccgaltgacccggtcctatccta 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2227 ATTTGACATCACTGGAAGTGTGTATATGTCGAGAAACAATTTGAAGGAAAGTTCCGC 2286
QY 500 gagaaactcacagttattcaagccttaagttgtagtctcagggaaatgattgtgtg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2287 AATGTTGGTAATATACGATGACCTTCACATTTTGTGATGTCATCAATATAGTTTCAGAG 2346
QY 560 gaacaattcc 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2347 GAGAGCTCCC 2356

```

## RESULT 15

```

VI4519
ID VI4519 standard; DNA; 3979 BP.
AC VI4519;
DE 20-MAY-1998 (first entry)
DE CF-5 pathogen resistance gene variant #2.
KW Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW tomato leaf mould; Phytophthora resistance; ss.
OS Lycopersicon pimpinellifolium.
FH Key Location/Qualifiers
FT CDS 653..3560
    /tag= a
    /product= CF-5 pathogen resistance gene
PN MO9743429-A1.
PD 20-NOV-1997.
PE 08-MAY-1997; G01249.
PR 24-SEP-1996; GB-019924.
PR 09-MAY-1996; GB-009681.
PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR P-PSDB: M41310.
PT Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT for production of transgenic plants resistant to pathogens e.g.
PT tomato leaf mould C. fulvum in tomatoes
PR Claim 6; Fig 1b; 75pp; English.
CC This sequence is an example of the polynucleotide of the invention, and
CC is able to confer pathogen resistance on a plant. It is one of two tomato
CC CF-5 gene variants, which offer resistance against the pathogen
CC Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC the gene into plant cells and regenerating plants from the cells;
CC asexually or sexually produced offspring can also be subsequently
CC produced. Expression of the gene in plant cells can confer pathogen
CC resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC Oligonucleotides with sequences complementary to the gene or fragments
CC of it, are useful in anti-sense techniques to reduce gene expression. The
CC nucleic acids/polynucleotides are useful as hybridisation probes to
CC identify other genes/fragments conferring pathogen resistance on plants
CC e.g. Phytophthora resistance in potatoes. Homologies between Cf-5 and
CC Cf-9 may be used to identify further resistance genes of this class.
SQ Sequence 3979 BP; 1218 A; 716 C; 698 G; 1347 T;

```

Query Match 7.0%; Score 62.8; DB 1; Length 3979;  
 Best Local Similarity 53.2%; Pred. No. 1.2e-07;  
 Matches 133; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

QY 320 aactctacaaaagagattcaaggaaactatatacctcttgagcttggaatctgaagagtc 379
    ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2107 ATCTTTACAAATAATACAGCTTTCGCTCTATTCCTGCTTATTCGCAATATAGAAATTC 2166
QY 380 taatcaagttggatctgttacacaacatctcacccgggaaatccatctctcttg9aa 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2167 TGCAAACTCTGTTTCACGATTAACGATCTCATTTGGGAAATTCCTTCATTGTGTGCA 2226
QY 440 aattgaagtcacttgcttcttgcggtctaacgaaacccgaltgacccggtcctatccta 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2227 ATTTGACATCACTGGAAGTGTGTATATGTCGAGAAACAATTTGAAGGAAAGTTCCGC 2286
QY 500 gagaaactcacagttattcaagccttaagttgtagtctcagggaaatgattgtgtg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2287 AATGTTGGTAATATACGATGACCTTCACATTTTGTGATGTCATCAATATAGTTTCAGAG 2346
QY 560 gaacaattcc 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2347 GAGAGCTCCC 2356

```

Search completed: June 23, 2000, 22:53:35  
 Job time: 40611 sec



4

5

6 7 8 9 10 11

12



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:45:40 ; Search time 274.21 Seconds

(without alignments)  
423.787 Million cell updates/sec

Title: US-09-180-798-28

Perfect score: 894

Sequence: 1 ggaccgattcaagctccga.....gtacttgattcttatia 894

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_MA:\*  
1: /cgn2\_6/prodata/2/1na/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/2/1na/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/2/1na/5C.COMB.seq:\*  
4: /cgn2\_6/prodata/2/1na/5D.COMB.seq:\*  
5: /cgn2\_6/prodata/2/1na/6.COMB.seq:\*  
6: /cgn2\_6/prodata/2/1na/PCRM.COMB.seq:\*  
7: /cgn2\_6/prodata/2/1na/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	62	6.9	1554	4	US-08-587-680A-24
2	61.2	6.8	3921	4	US-08-567-375-3
3	61.2	6.8	3921	4	US-08-587-680A-3
4	61.2	6.8	3921	4	US-08-475-891A-3
5	59.6	6.7	6256	3	US-08-475-891A-1
6	59.6	6.7	6256	4	US-08-567-375-1
7	59.6	6.7	6256	4	US-08-587-680A-1
8	58.8	6.6	831	4	US-08-567-375-15
9	53.8	6.0	2880	4	US-08-666-271-4
10	53.8	6.0	3905	4	US-08-666-271-1
11	52.4	5.9	3733	3	US-08-473-553A-1
12	40.6	4.5	2075	1	US-08-238-163-3
13	36	4.0	1116	2	US-08-244-646-14
14	35.6	4.0	1380	5	US-08-968-563-5
15	35.4	3.9	3095	7	5231168-1
16	35	3.9	19124	4	US-08-487-826B-13
17	34.4	3.8	1058	1	US-08-238-163-1
18	34.4	3.8	2917	2	US-08-592-936B-20
19	34.4	3.8	2917	4	US-09-111-573-20
20	34.4	3.8	7218	1	US-08-332-463-14
21	33.8	3.8	1868	4	US-08-960-022-5
22	33.6	3.8	3440	2	US-08-471-791-27
23	33.6	3.8	3440	6	PCT-US81-01746-27
24	32.8	3.7	792	2	US-08-244-646-16
25	32.8	3.7	2210	1	US-07-710-361-2
26	32.2	3.6	4599	2	US-08-431-080-27
27	32.2	3.6	4599	4	US-08-938-534-27

28	32.2	3.6	4951	4	US-08-867-030B-5	Sequence 5, Appl1
29	32.2	3.6	4951	6	PCT-US95-06119-5	Sequence 5, Appl1
30	31.2	3.5	2241	6	PCT-US95-10509-1	Sequence 1, Appl1
31	31	3.5	426	2	US-08-442-063A-35	Sequence 35, Appl1
32	31	3.5	567	2	US-08-442-063A-38	Sequence 38, Appl1
33	31	3.5	711	2	US-08-442-063A-41	Sequence 41, Appl1
34	31	3.5	849	2	US-08-442-063A-44	Sequence 44, Appl1
35	31	3.5	909	2	US-08-347-471-3	Sequence 3, Appl1
36	31	3.5	924	2	US-08-442-063A-47	Sequence 47, Appl1
37	31	3.5	1002	2	US-08-442-063A-26	Sequence 26, Appl1
38	31	3.5	1026	2	US-08-272-919-1	Sequence 1, Appl1
39	31	3.5	1026	6	US-08-619-916-1	Sequence 1, Appl1
40	31	3.5	1026	6	PCT-US95-08542-1	Sequence 1, Appl1
41	31	3.5	1593	7	5340934-3	Patent No. 5340934
42	30.8	3.4	2674	4	US-08-926-724-2	Sequence 2, Appl1
43	30.6	3.4	2044	1	US-08-071-601-3	Sequence 3, Appl1
44	30.6	3.4	2044	7	US-08-621-100-3	Sequence 3, Appl1
45	30.6	3.4	11225	7	5182210-9	Patent No. 5182210

## ALIGNMENTS

RESULT 1  
US-08-587-680A-24  
Sequence 24, Application US/08587680A  
Patent No. 5977434  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Confering  
TITLE OF INVENTION: Disease Resistance in Plants  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/587, 680A  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373, 375  
FILING DATE: 17-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/475, 891  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004, 645  
FILING DATE: 29-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/567, 375  
FILING DATE: 04-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-058940US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs







```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1..2676, 3520..3918)
; OTHER INFORMATION: /product= "Xa-21"
;
US-08-587-680A-3

Query Match          6.8%; Score 61.2; DB 4; Length 3921;
Best Local Similarity 52.3%; Pred. No. 1.1e-08;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 313 tataggaactcacaaaacgagattcaaggaactatccttcgagcttggaaatctg 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1138 TTCTTGCACTGATTAATGATAGATCAGAGACGATTCGAGAGATTTGGCAATCTT 1197
QY 373 aagaagctaatcagcttggatctgtacaaacaatctcaccgggaaatcccatctct 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1198 ATGGGCTTACACATCTCATCTCTGCAACAACAATTTCAGAGGGCTCTTCCATCATCG 1257
QY 433 ttgggaattgaagtcactgttttttgcggttaacgaaacggattgacggctct 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1258 TTGGGAGGCTTAATAACTTAGGCATTTCTACCTGCGCTACGAAACAACCTTGAGCGGTTGCG 1317
QY 493 atccctagagaactcacagttatttcaagccttaaggttgatgctcagggatgat 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1318 ATCCCGTTGGCCATAGGAATCTTACTGAACCTTAATATCTTACTGCTCGGACACCAACAA 1377
QY 553 ttgttggaaacaattcca 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1378 TTCAGTGTGTGATACCA 1395

RESULT 4
US-08-475-891A-3
; Sequence 3, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5992 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(512..3149, 3993..4393)
; OTHER INFORMATION: /product= "RRK-B"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-B from rice (Oryza
; OTHER INFORMATION: sativa)"
;
US-08-475-891A-3

Query Match          6.8%; Score 61.2; DB 3; Length 5992;
Best Local Similarity 52.3%; Pred. No. 1.4e-08;
Matches 135; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 313 tataggaactcacaaaacgagattcaaggaactataccttcgagcttggaaatctg 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1649 TTCTTGCACTGATTAATGATAGATCAGAGACGATTCGAGAGATTTGGCAATCTT 1708
QY 373 aagaagctaatcagcttggatctgtacaaacaatctcaccgggaaatcccatctct 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1709 ATGGGCTTACACATCTCATCTCTGCAACAACAATTTCAGAGGGCTCTTCCATCATCG 1768
QY 433 ttgggaattgaagtcactgttttttgcggttaacgaaacggattgacggctct 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1769 TTGGGAGGCTTAATAACTTAGGCATTTCTACCTGCGCTACGAAACAACCTTGAGCGGTTGCG 1828
QY 493 atccctagagaactcacagttatttcaagccttaaggttgatgctcagggatgat 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1829 ATCCCGTTGGCCATAGGAATCTTACTGAACCTTAATATCTTACTGCTCGGACACCAACAA 1888
QY 553 ttgttggaaacaattcca 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1889 TTCAGTGTGTGATACCA 1906

RESULT 5
US-08-475-891A-1
; Sequence 1, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
```



```

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-05891005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice (Oryza
; OTHER INFORMATION: sativa)"
;
; US-08-475-891A-1

```

```

Query Match 6.7%; Score 59.6; DB 3; Length 6256;
Best Local Similarity 51.9%; Pred. No. 4.2e-08;
Matches 134; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 313 tatagtgaacctctcaaaaagagatccaaggaactataacctcttgctgctgaaatctg 372
DB 2806 TTCTTGACCTTGATTGAAATGAAGATCAAGAGAAAGCAATTCGCAATTTGGCAATCTT 2865

QY 373 aagagctcaatcaagttgagctgtacaacaacaatctcacccgggaaatcccatctctc 432
DB 2866 ATTGGCTTACAACATCTCTATCTCTGCAACAACATTTCAAGAGGCTACCTTCATCATCG 2925

QY 433 ttgggaaaattggaagctactgttttttgcgcgttaagcaaaaacgattgacgcgtctc 492
DB 2926 TTGGGACAGGCTTAGAAACTTAGGCATTTCTAGTCGCTACGAAACAACTTGAGGCGTTG 2985

QY 493 attccatagaaactcaagctatttcaagcttaagttgtgtgctcaaggaaatgat 552
DB 2986 ATCCCATTTGGCCATGAGAAATCTTACGAACTTAATATCTTACTGCTGCGCACCAACAA 3045

QY 553 ttgtgtggaacaattcca 570
DB 3046 TTCAGTGGTTGATACCA 3063

```

```

RESULT 6
US-08-567-375-1
; Sequence 1, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.

```

```

; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1648..4383, 5178..5513)
; OTHER INFORMATION: /product= "RRK-F"
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance gene RRK-F from rice
; OTHER INFORMATION: (Oryza sativa)"
;
; US-08-567-375-1

```

```

Query Match 6.7%; Score 59.6; DB 4; Length 6256;
Best Local Similarity 51.9%; Pred. No. 4.2e-08;
Matches 134; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 313 tatagtgaacctctcaaaaagagatccaaggaactataacctcttgctgctgaaatctg 372
DB 2806 TTCTTGACCTTGATTGAAATGAAGATCAAGAGAAAGCAATTCGCAATTTGGCAATCTT 2865

QY 373 aagagctcaatcaagttgagctgtacaacaacaatctcacccgggaaatcccatctctc 432
DB 2866 ATTGGCTTACAACATCTCTATCTCTGCAACAACATTTCAAGAGGCTACCTTCATCATCG 2925

QY 433 ttgggaaaattggaagctactgttttttgcgcgttaagcaaaaacgattgacgcgtctc 492
DB 2926 TTGGGACAGGCTTAGAAACTTAGGCATTTCTAGTCGCTACGAAACAACTTGAGGCGTTG 2985

QY 493 attccatagaaactcaagctatttcaagcttaagttgtgtgctcaaggaaatgat 552
DB 2986 ATCCCATTTGGCCATGAGAAATCTTACGAACTTAATATCTTACTGCTGCGCACCAACAA 3045

```















```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,163
; FILING DATE: 03-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 421..1401
; US-08-238-163-3

```

```

Query Match          4.5%; Score 40.6; DB 1; Length 2075;
Best Local Similarity 48.5%; Pred. No. 0.01;
Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

```

```

QY 343 ggaactatactcttgagcttggaatctgaagatcctaactcagtttgatctgtaaac 402
DB 733 ggaacattccactgacattgacgagctcaaaatctcaaaatctgaagctcagcttc 792
QY 403 aacaaatccaccgggaaataccacatctcttggaagaattgaaatcagtttttttg 462
DB 793 actaaccttaccaggtccgacatccctgaattccttactgacgtgaagaattgacgttc 852
QY 463 cggcttaacgaacacgattgacggttcctattcctctagaagaactcagttatttcaagc 522
DB 853 gattggaattacacatatttaccggaacatccctcttccctctccagcttccggaat 912
QY 523 cctaagttggtgatgtctcaggaagtattgtgtggaacaattcagta 573
DB 913 ttgcraagcattgactttagatctgtaacaaactcaccggaacaaatccgga 963

```

```

RESULT 13
US-08-244-646-14
; Sequence 14, Application US/08244646
; Patent No. 5744692
; GENERAL INFORMATION:
; APPLICANT: Cervone, Felice
; APPLICANT: De Lorenzo, Giulia
; APPLICANT: Salvi, Giovanni
; APPLICANT: Albersheim, Peter
; APPLICANT: Darvill, Alan
; APPLICANT: Bergmann, Carl
; TITLE OF INVENTION: Nucleotide Sequences Coding An
; TITLE OF INVENTION: Endopolygalacturonase Inhibitor
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sally A. Sullivan
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,646
; FILING DATE: 06-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM 91A 000915
; FILING DATE: 06-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/IT/00158
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sullivan, Sally A.
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 19-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Phaseolus vulgaris
; STRAIN: Saxa
; IMMEDIATE SOURCE:
; CLONE: lambda PGIP-3.3
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1026
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1027..1116
; US-08-244-646-14

```

```

Query Match          4.0%; Score 36; DB 2; Length 1116;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 90; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```

```

QY 319 gaactcaaaaagagattcaaggaactataactcttgagcttggaatctgaagagt 378
DB 835 gatctgaggaaacacgattatgagacgctacaggacgtaacgacgttaagttt 894
QY 379 ctataagttgattctgacacacaaactcaccgggaaatccatcttcttgga 438
DB 895 ctgcagaatttaatttgacttcaacaaatctgctcggaatcttccaaagtggaac 954
QY 439 aaattgaagctactgttttttgcggttaacgaacacgattgacggttcattcct 498
DB 955 ttgaaaggttgacgtttcttctttagcacaacaaagtgctgtggttcttctctt 1014

```

```

RESULT 14
US-08-968-563-5
; Sequence 5, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATEWAY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY M. WHITED
; APPLICANT: VASANTHA NAGARAJAN

```



```

; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 4 CAMBRIDGE PLACE
; STREET: 1870 SOUTH WINTON ROAD
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,563
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: NOVEMBER 13, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9982
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1380 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: GPD1
; US-08-968-563-5

```

```

Query Match          4.0%; Score 35.6; DB 5; Length 1380;
Best Local Similarity 58.5%; Pred. No. 0.25;
Matches 62; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

QY 449 cactgttttttgcggttaagcaaaacgattgacgggtctctattcctagaagactca 508
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 479 CACATCATTTTTCGCCCGTATCTGTAGCCATGGAAGGTCAATGTCATGTCACACGTCA 538
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 509 cagtatttcaggcccttaagttgtatgtctcagggaatgattt 554
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 539 GAGCTATCTCCTGTCTAAGGGTTTGAAGTTGAGTGTGCTAAGGTCGT 584
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
5231168-1/c
; Patent No. 5231168
; APPLICANT: DIEGIEL, MORTEN,BORRE, MARTIN,JEPSEN, SOREN,
; VUUST, JENS,RIENCK, KLAUS,WIND, ANNETTE,JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989

```

```

; SEQ ID NO:1:
; LENGTH: 3095
5231168-1

```

```

Query Match          4.0%; Score 35.4; DB 7; Length 3095;
Best Local Similarity 53.2%; Pred. No. 0.43;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

```

```

QY 741 tgtttgtaatactatataagaccttggttcacgtatataatgaagcttgagagacagt 800
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2495 TATTTCTATATATACAGAAATCCTAGTATCTTTATATATATATATATATATATTA 2436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 801 aacttgcaatgataatgtatgtatgtatgtagaaaagtgaatgagaattgcttgtaattga 860
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2435 ATAAATATTAGTTCATATATATAGAAAAATTTATATATATTTTATATATTTT 2376
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 861 ttgtgttcttatgtaactt 881
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2375 TTTTTTTTTTTTTTGTAAAT 2355
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Search completed: June 23, 2000, 22:46:00
Job time: 40306 sec

```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 19:06:38 ; Search time 6198.48 Seconds  
(without alignments)  
584.592 Million cell updates/sec

Title: US-09-180-798-28  
Perfect score: 894  
Sequence: 1 ggaacgattcaagctccga.....gttaactgattcttattta 894

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: em\_est10:\*  
11: em\_est11:\*  
12: em\_est12:\*  
13: em\_est13:\*  
14: em\_est14:\*  
15: em\_est15:\*  
16: em\_est16:\*  
17: em\_est17:\*  
18: em\_est18:\*  
19: em\_est19:\*  
20: gb\_est1:\*  
21: gb\_est2:\*  
22: gb\_est3:\*  
23: gb\_est4:\*  
24: gb\_est5:\*  
25: gb\_est6:\*  
26: gb\_est7:\*  
27: gb\_est8:\*  
28: gb\_est9:\*  
29: gb\_est10:\*  
30: gb\_est11:\*  
31: gb\_est12:\*  
32: gb\_est13:\*  
33: gb\_est14:\*  
34: gb\_est15:\*  
35: gb\_est16:\*  
36: gb\_est17:\*  
37: gb\_est18:\*  
38: gb\_est19:\*  
39: gb\_est20:\*  
40: gb\_est21:\*  
41: gb\_est22:\*  
42: gb\_est23:\*  
43: gb\_est24:\*  
44: gb\_est25:\*

45: gb\_est26:\*  
46: gb\_est27:\*  
47: gb\_est28:\*  
48: gb\_est29:\*  
49: gb\_est30:\*  
50: gb\_est31:\*  
51: gb\_est32:\*  
52: em\_est20:\*  
53: em\_est21:\*  
54: em\_est22:\*  
55: em\_est23:\*  
56: em\_est24:\*  
57: em\_est25:\*  
58: em\_est26:\*  
59: gb\_est33:\*  
60: gb\_est34:\*  
61: gb\_est35:\*  
62: gb\_est36:\*  
63: gb\_est37:\*  
64: gb\_est38:\*  
65: em\_est27:\*  
66: em\_est28:\*  
67: em\_est29:\*  
68: em\_est30:\*  
69: gb\_est39:\*  
70: gb\_est40:\*  
71: gb\_est41:\*  
72: gb\_est42:\*  
73: gb\_est43:\*  
74: gb\_est44:\*  
75: em\_est31:\*  
76: em\_est32:\*  
77: em\_est33:\*  
78: em\_est34:\*  
79: gb\_est45:\*  
80: gb\_est46:\*  
81: gb\_est47:\*  
82: gb\_gss1:\*  
83: gb\_gss2:\*  
84: gb\_gss3:\*  
85: gb\_gss4:\*  
86: em\_gss1:\*  
87: em\_gss2:\*  
88: em\_gss3:\*  
89: em\_gss4:\*  
90: gb\_gss5:\*  
91: gb\_gss6:\*  
92: gb\_gss7:\*  
93: gb\_gss8:\*  
94: gb\_gss9:\*  
95: em\_gss5:\*  
96: em\_gss6:\*  
97: em\_gss7:\*  
98: em\_gss8:\*  
99: em\_gss9:\*  
100: em\_gss10:\*  
101: em\_gss11:\*  
102: gb\_gss10:\*  
103: gb\_gss11:\*  
104: em\_gss12:\*  
105: gb\_gss12:\*  
106: gb\_gss13:\*  
107: gb\_gss14:\*  
108: gb\_gss15:\*  
109: gb\_gss16:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES



No.	Score	Match	Length	DB	ID	Description
1	419.8	47.0	450	20	T04109	T04109 59 Lambda-P
2	387.4	43.3	443	63	A1999615	A1999615 70156713
3	265	29.6	556	23	R89988	R89988 16353 Lambda
4	256.4	28.7	353	23	H37300	H37300 15429 Lambda
5	223.8	25.0	534	33	A4394359	A4394359 25942 Lam
6	220.6	24.7	608	81	AW443205	AW443205 EST308135
7	219.6	24.6	720	64	AM031110	AM031110 EST274417
8	213.8	23.9	430	79	AM0307218	AM0307218 sfs4c07.Y
9	210.4	23.5	657	80	AM350720	AM350720 GM210009A
10	209.6	23.4	570	50	A1676939	A1676939 605047407
11	208.6	23.3	541	64	AM036855	AM036855 614019631
12	207.8	23.2	637	80	AM350549	AM350549 GM210009A
13	205	22.9	646	74	AM021939	AM021939 EST287950
14	204.2	22.8	545	64	AM037836	AM037836 EST279465
15	202.6	22.7	305	25	N65416	N65416 20456 Lambda
16	199.2	22.3	511	74	AM221278	AM221278 EST297747
17	197.6	22.1	714	64	AM030188	AM030188 EST273443
18	196.8	22.0	364	42	A1100683	A1100683 33804 Lam
19	194.8	21.8	579	43	AU031292	AU031292 AU031292
20	189.4	21.2	417	21	D46231	D46231 R1C810770A
21	186.6	20.9	616	42	AF074734	AF074734 AF074734
22	185.2	20.7	348	42	A1100682	A1100682 33803 Lam
23	184	20.6	695	64	AM038166	AM038166 EST279825
24	183.4	20.5	460	35	C72525	C72525 C72525 R1C6
25	182.2	20.4	562	50	A1691407	A1691407 606015E09
26	182.2	20.4	599	74	AM219797	AM219797 EST302279
27	173.2	19.4	334	23	H37195	H37195 15324 Lambda
28	171.8	19.2	599	63	AM011134	AM011134 SP17B03 P
29	171	19.1	339	42	A1100679	A1100679 33800 Lam
30	168.2	18.8	469	63	A1992868	A1992868 701493826
31	167	18.7	676	74	AM220075	AM220075 EST302558
32	163	18.2	569	59	A1775448	A1775448 EST26548
33	156.6	17.5	423	35	C22371	C22371 C22371 R1C6
34	155.6	17.4	336	23	H37296	H37296 15425 Lambda
35	154	17.2	401	37	AA712221	AA712221 31949 Lam
36	151.6	17.0	612	47	A1496325	A1496325 sb05c09.Y
37	149.4	16.7	458	59	A1777982	A1777982 EST258861
38	149.4	16.7	463	74	AM221200	AM221200 EST297669
39	149.2	16.7	502	79	AM286582	AM286582 L61_334_F
40	148.8	16.6	514	35	C22372	C22372 C22372 R1C6
41	147.8	16.5	368	64	AM032898	AM032898 EST276457
42	147.8	16.5	423	59	A1774881	A1774881 EST255981
43	147.8	16.5	423	62	A1896737	A1896737 EST266180
44	147.8	16.5	488	64	AM031917	AM031917 EST275371
45	147.8	16.5	497	59	A1772079	A1772079 EST253179

## ALIGNMENTS

RESULT	1	450 bp	mRNA	EST	06-NOV-1997
LOCUS	T04109	59 Lambda-PRL1	Arabidopsis thaliana	CDNA clone	SCH97P, mRNA
DEFINITION	sequence.				
ACCESSION	T04109				
VERSION	EST				
KEYWORDS	T04109.1	GI:315269			
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1	(bases 1 to 450)			
AUTHORS	Neman, F., deStrufin, F.J., Green, P., Keegstra, K., Kende, H., McIltoosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E., and Somerville, C.				
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones				
JOURNAL	Plant Physiol.	106,	1241-1255	(1994)	

MEDLINE	95148729.
COMMENT	Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel: 517-353-0854 Fax: 517-353-9168 Email: 22313cne@bm.ci.msu.edu.
FEATURES	Location/Qualifiers
SOURCE	1..450 /organism="Arabidopsis thaliana" /strain="var columbia" /db_xref="taxon:3702" /clone="SCH97P" /note="Vector: lambda Shlox-1; Site_1: EcoRI; Site_2: HindIII; lambda PRL1 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The library was made in Novagen's lambda Shlox-1 with (oligo dt primed) directional inserts cloned between the EcoRI and HindIII sites."
BASE COUNT	125 a 94 c 81 g 147 t 3 others
ORIGIN	
Query Match	47.0%; Score 419.8; DB 20; Length 450;
Best Local Similarity	98.4%; Pred. No. 3.6e-92;
Matches 443; Conservative	0; Mismatches 5; Indels 2; Gaps 2;
QY	24 ggaagctcttcacgcgcttcgcgcgagcttcacagatccagaaatgtttcagaagt 83
DB	1 GCAGGCTCTTACGCGCTTCGCGGAGCTTACATCCAGCAATGTTGTCAAGTGG 60
QY	84 ggaacacactctgttaacactctgtactgtgttcacgtcactgttaacaacacatca 143
DB	61 GGATCCACTCTGTTAATCTTGATCTTGATCTTGATCTTGATCTTGATCTTGATCT 120
QY	144 agtcaactctgtcgttgatcttggaatcaactatctcgacatctagtaactgaaacttg 203
DB	121 AGTCACTCGTCTGATTTGGGGAATCAACTTACTGTGACATCTAGTACCTGAACTGG 180
QY	204 gaagcttgaaacttcaatctatctgtagtaatactacactcttgcccttttgatgact 263
DB	181 GAACTTGAACTTCAATATCTGATGAAATCACTCTTTGCCCTTGATTAATCT 240
QY	264 gaaacattacattacacacacataataacatttgcttgagtcataatagtaact 323
DB	241 GAAACATTACATTACATGATCAATATTAATTTGCTTGATGATATAGGAAGCT 300
QY	324 ctacaaaacagagatcaaggaactacactctcgagcttggaatctgaagatctaat 383
DB	301 CTACAAAACAGATCAAGGAATCTATACCTTGAGCTTGAATCTGAAATCTGAAATCTAT 360
QY	384 caatttgatctgtacaaacacatctacac--gggaatcccatctcttc--gggaaa 441
DB	361 CAATTGATCTGTACACACATCTCACCGGGAATCCCATCTCTTGGGGAAAA 420
QY	442 ttgaagtaactgttttttcgagcttaac 471
DB	421 TTGAAGTACTGTGTTTTTGGCGGNTTAC 450
RESULT	2
LOCUS	A1999615 443 bp mRNA EST 08-SEP-1999
DEFINITION	70156713 A. thaliana, Columbia Col-0, rosette-3 Arabidopsis thaliana cDNA clone 70156713, mRNA sequence.
ACCESSION	A1999615



VERSION A1999615.1 GI:5846520  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II: Brassicales: Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 443)  
 AUTHORS

Chen, J., Montyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, D., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Monanouthana, M., Nguyen, C., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.  
 Arabidopsis thaliana Gene Expression Microarray  
 Unpublished (1999)  
 On Mar 10, 1998 this sequence version replaced gi:2948106.  
 Contact: David Smoller, Ph.D.  
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.  
 4633 World Parkway Circle, St. Louis, MO 63134, USA  
 Tel: 877-577-2733  
 Fax: 314-427-3324  
 Email: service@genomesystems.com.

FEATURES  
 SOURCE

1..443  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia Col-0"  
 /db\_xref="taxon:3702"  
 /clone="70156713"  
 /clone\_1lb="A. thaliana, Columbia Col-0, rosette-3"  
 /tissue\_type="rosette"  
 /dev\_stage="4 - 7 weeks"  
 /note="Vector: pSPORT; Site\_1: NotI; Site\_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."  
 BASE COUNT 135 a 88 c 54 g 124 t 42 others  
 ORIGIN

Query Match 43.3%; Score 387.4; DB 63; Length 443;  
 Best Local Similarity 90.2%; Pred. No. 2.6e-84;  
 Matches 388; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

455 ttttttcggttaacgaacacgattgacggttcctattcctagagaaactcacagttta 514  
 |||||  
 Db 430 TTTTTCGCGCTNACGAAAACNAGATNGACNGTNTATTCCTGAGNACCTCAGATT 371  
 515 ttccaagccttaagttgttgcgttcacaggaatgattgtgtgtaacaactccagtag 574  
 |||||  
 Db 370 TTTCAGCCNTAAANTNGTTCATGCTCAGNAGATGATNGTGTGNAACAATCCAGTAG 311  
 575 aaggacctttgaacacactcctatgacaaactttgagaaactcagatcagattgagagac 634  
 |||||  
 Db 310 AAGGACNTTTGNNAACATTCCTGATNGCNAACCTTGAACNACCACTGATGAGAGGAC 251  
 635 cagaactactaggtcttgcgagctatgacacaaacttgcaacttaaaagaagtgaagaac 694  
 |||||  
 Db 250 CGAAGCTACTAGGCTTCGAGNTWTGACACCAATTGCACTTNAAGAAAGTTGAAGAAC 191  
 695 ctataaagaagaattgagtgatgagccttgaagaactcgttgcgaagtttgtaaatccta 754  
 |||||  
 Db 190 CTNTAANGAAGATGTTAGGTGACNTTGTAAAGANTCTGTCAGCAAGTNTTTGTAATCTCA 131

QY 755 tataagccttgtttcattatataatgaagcctttgagacagtaactgcacattat 814  
 |||||  
 Db 130 TATAGACNTGTTCATGTTATATGAAACCTTGAGAGCACTAATNTGCATGAT 71

QY 815 tggatgtgagaaagtgaatgaatgagattccttttaattgattgtttcctat 874  
 |||||  
 Db 70 TNNTANTNGTAGNANAAGTTGNNAAGATTCCTTTGTAATGATNTGTCCTTTCTTAT 11

QY 875 gtaactgaa 884  
 |||||  
 Db 10 GTAANTGAA 1

RESULT 3

R89998 556 bp mRNA EST 30-DEC-1997  
 LOCUS 16353 Lambda-PRL2 Arabidopsis thaliana cDNA clone 187H577, mRNA  
 DEFINITION  
 ACCESSION R89998  
 VERSION R89998.1 GI:957538

KEYWORDS  
 SOURCE

thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II: Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 556)

NEWMAN, J., deBRUIJN, F. J., GREEN, P., KEESSTRA, K., KENDE, H., MCINTOSH, L., ONLROGGE, J., RAIKHEL, N., SOMERVILLE, S., THOMASHOW, M., RETZEL, E. and SOMERVILLE, C.

Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)

JOURNAL MEDLINE

95148729  
 On Apr 14, 1993 this sequence version replaced gi:693023.

COMMENT

Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 22313tch@dm.cl.msu.edu  
 Seq primer: T7 dye primer.  
 Location/Qualifiers

FEATURES  
 source

1..556  
 /organism="Arabidopsis thaliana"  
 /strain="var. Columbia"  
 /db\_xref="taxon:3702"  
 /clone="187H577"  
 /clone\_1lb="lambda-PRL2"  
 /note="Vector: lambda Z1p-lox; Site\_1: SalI; Site\_2: NotI; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark - rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Z1p-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."  
 BASE COUNT 151 a 118 c 123 g 146 t 18 others  
 ORIGIN

Query Match 29.6%; Score 265; DB 23; Length 556;  
 Best Local Similarity 80.0%; Pred. No. 1.6e-54;  
 Matches 335; Conservative 0; Mismatches 79; Indels 5; Gaps 2;

QY 313 tatagtgaaccttcaaaacgagattcgaagacatataccttctgagttggaattcgt 372  
 |||||  
 Db 49 TATCTAGAGCTCTCAAAAACACATCCAGAGACTATACCTTCGAACTTGGAATCTTG 108











DEFINITION EST274417 tomato callus, TAMU Lycopersicon esculentum cDNA clone  
 CLEC13C21 similar to leucine-rich repeat protein (LRR), mRNA  
 sequence.  
 ACCESSION AM031110  
 VERSION AM031110.1 GI:5889866  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Lycopersicon esculentum  
 tomato.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 REFERENCE 1 (bases 1 to 720)  
 AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Viston,T.,  
 Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,  
 Kohnmug,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and  
 Giovannoni,J.  
 TITLE Generation of ESTs from tomato callus tissue  
 JOURNAL Unpublished (1999)  
 COMMENT On Dec 20, 1995 this sequence version replaced gi:1133546.  
 CONTACT: David Frisch  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 4366  
 Fax: 864 656 4293  
 Email: dfrisch@CLEMSON.EDU  
 5 prime sequence.  
 FEATURES  
 source  
 1. /720  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEC13C21"  
 /clone\_1lb="tomato callus, TAMU"  
 /tissue\_type="callus"  
 /dev\_stage="25-40 days old"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: pBluescript SK(-); Site,1: EcoRI; Site,2:  
 XhoI; supplier: Giovannoni laboratory; cDNA - Cotyledons  
 of seedlings 7-10 days post-germination were excised, cut  
 at both ends and placed on MS medium with no selection.  
 Mixed callus was harvested at 25 and 40 days and included  
 undifferentiated masses. Tomato Callus EST Library"  
 BASE COUNT 200 a 149 c 143 g 228 t  
 ORIGIN  
 Query Match 24.6%; Score 219.6; DB 64; Length 720;  
 Best Local Similarity 64.4%; Pred. NO. 1.8e-43;  
 Matches 397; Conservative 0; Mismatches 129; Indels 90; Gaps 1;

Db 408 --GAGCTTACAAATATATATACAGGAACCATCCCTAAGAGCTCGCTACTTGAAG 464  
 Qy 376 agtctaatcagtttgatcgtctacaaacaatcaccgagggaatccattcttgg 435  
 Db 465 AGCCTTATAGTCTGATCTGTACACACAAATATTTCCGGGACAAATCTCTACTTACTT 524  
 Qy 436 ggaatattgaagtcactgttttttgcggcttaacgaacagattgacggtccatt 495  
 Db 525 GGAACCTGAAAAACCTTGTTTCTTCGCTAAATATACAAAGCTACAGACCAATC 584  
 Qy 496 cctagagactcacagattattcctaagccttaagttgtatgctcgaaggaaatgg 555  
 Db 585 CCAAGAGAACTTACTACACTTCTAGCCTGAAGTTGTGATGTCCTGAAATATGATTG 644  
 Qy 556 tgtggaacaattccaagagaagacgttttgaaacacttccttgcaaaacttgaagc 615  
 Db 645 TGTGAAACAATTCCTACTTCTGTCTGATTTGAGCATATCTCTAAACAATTCAGACAC 704  
 Qy 616 aacctgagattgagag 631  
 Db 705 AATCTGCACCTTGAAG 720

RESULT 8  
 AM307218 430 bp mRNA EST 20-JAN-2000  
 LOCUS s54c07.y1 Gm-cl009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-cl009-3781 5' similar to TR:096477 Q96477 LRR PROTEIN. ; mRNA  
 sequence.

ACCESSION AM307218  
 VERSION AM307218.1 GI:6719571  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 Glycine.  
 1 (bases 1 to 430)  
 Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V.,  
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theisling,B., Allen,M.,  
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,  
 Schurr,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterson,R. and Wilson,R.

TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT On Dec 20, 1995 this sequence version replaced gi:1133696.  
 CONTACT: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@wustl.wustl.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134 For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or  
 info@genomesystems.com web site: www.genomesystems.com  
 Seq primer: -40RP from GIDCO  
 High quality sequence stop: 426.

#### FEATURES

source  
 1. /430  
 Location/Qualifiers  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl009-3781"  
 /clone\_1lb="Gm-cl009"  
 /lab\_host="XLI0-Gold"  
 /note="Vector: pBluescript II XR; Site,1: EcoRI; Site,2:  
 XhoI; The mRNA was isolated from entire roots of  
 2-month-old 'Williams' plants that were greenhouse grown







Query Match	23.4%	Score 209.6	DB 50	Length 570
Best Local Similarity	73.9%	Pred. No. 4,8e-41		
Matches	266	Conservative	0	Mismatches 94; Indels 0; Gaps 0;
QY	313	tatagtagactctacaaaacgagatctcaaggaactatccctctgagcttggaaatctg	372	
Db	541	TATCGGACCTTTAAAGATATACACTTCAGAGAAACAATCCCTTCGAGCTTGCGAACTTG	482	
QY	373	aagagctcaaatagtttgatctgcaacacaaacatctacccgggaaaatcccatctct	432	
Db	481	AGAAATCTAATAGCTTCGACTTGTGACGAACACATTTGGGGACAACTTCCTCGCC	422	

[illegible]



```

Db 155 CTTGGAAGTTGAGTCCCTTGATCTTCTGCTCAAGGCAATATTGACTGGGCC 214
OY 493 attctagaagaccacagattattccaagccttaaaagtgtgatgtctcaaggaaatgat 552
Db 215 ATCCCAAGGGAACACGCTGGAATATCTAGTCTCAAGTGTGATGTTTCAAGTATGAT 274
OY 553 ttgtgtgaacaactccagtagaagaccccttgaacacatctcctatgcaaaatttgag 612
Db 275 TTGTGCGGGAGAGATTCTCCAGCTCTGACCATTTGAGCATTTCTCTGAGCACTTTGAG 334
OY 613 aacacacgtgaattggagagaccagactactagctgtctgcagcctatcaccaattgac 672
Db 335 AAGAAACCGCGCTTGGAAGGTCACAGACTACAGGCTGCGCATATACGACCACTTCGC 394

```

## RESULT 12

AM350549

```

LOCUS GM210009A10A12R Gm-r1021 Glycine max cDNA clone Gm-r1021-3095 3',
DEFINITION

```

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA

AM350549 637 bp mRNA



Email: dfrisch@CLEMSON.EDU  
5 prime sequence.  
Location/Qualifiers

## FEATURES

1..646

/organism="Lycopersicon esculentum"

/cultivar="74496"

/db\_xref="taxon:4081"

/clone\_id="CLENM3"

/tissue\_type="pericarp"

/dev\_stage="red ripe (7-20 days post-breaker)"

/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 181 a 126 c 138 g 201 t  
ORIGIN

Query Match 22.9%; Score 205; DB 74; Length 646;  
Best Local Similarity 64.6%; Pred. No. 6.3e-40;  
Matches 374; Conservative 0; Mismatches 115; Indels 90; Gaps 1;

QY 16 tcgaaggagacgctcttcacgcgcttcgcgagctatcagatccagaaatgttgtt 75  
DB 158 TCAGAGGAGGAGGAGCTTTGACGCCCTCCGCGAGCTATCTGACCCGGTAAAGTGT 217  
QY 76 cagagcttgatccacaccttgtaatactctgtactctgtctgtatctgtatctgt 135  
DB 218 CGAGCTGGAGTCCAAATCTGTATACCTGTGACCTGCTTCAGTCAATGCTCAACGGA 277  
QY 136 caccatcaagtcacgctcgtctgatttgggaattcaactatctgacatctagtaact 195  
DB 278 GATATCAAGTACTCGTGTGATCTTGGAACTCAAGTATATCGTCAATTTGTACT 337  
QY 196 gaacttgggaagcttgacaataatctgtatgtgaatcaatcaactctttgccttt 255  
DB 338 GAGCTCGGAAACCTTGAACTATCTACGATCTG----- 371  
QY 256 gattatctgaacaattacattacatcagtcacacataataattgttgatcatat 315  
DB 371 ----- 371  
QY 316 agtgaactctacaaaacgagattcaaggaactatcctctgagcttggaaatctgaag 375  
DB 371 ---GAGCTTTACAAAATAATATTCAGGAAACCATCCCTAAGGAGCTCGTAACTTGAG 427  
QY 376 agtcaaatcagcttgatctgtacacaaacatctcacgcggaaatccatcttcttg 435  
DB 428 AGCTTATTAAGTCTGATCTGTACAAACAAATATTTGGGACAAATCTCTACTTCACTT 487  
QY 436 ggaataatgaagcttcttcttgcggtctaaagaaacggaatgacgcgtctactt 495  
DB 488 GGAAGAACTGAAGAACTTTCTTTCGCTTAATGATTAACAACTAAGACGACCAATC 547  
QY 496 cctagagaactcacagcttatctcaagccttaagatgtgtatgtctcaggaatgattg 555  
DB 548 CCAAGAGAACTTACATGCTTTCTAGCCGAAAGTTGGAGATGCTCAATTAATGATTGG 607  
QY 556 tctggaacaattccagtagaagacatttggacaatt 594  
DB 608 TGTGGAACAATTCCTACTCTGTCTCATTTTGACATATT 646

RESULT 14  
AM037836 545 bp mRNA EST 18-OCT-1999  
LOCUS AM037836  
DEFINITION EST729465 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

ACCESSION AM037836  
VERSION AM037836.1 GI:5896590  
KEYWORDS EST.  
SOURCE  
ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 545)

D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J., Roming, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Niernm, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and Giovannoni, J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

On Jun 22, 1998 this sequence version replaced gi:3246851.

CONTACT: David Frisch

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Fax: 864 656 4293

Email: dfrisch@CLEMSON.EDU

5 prime sequence.

Location/Qualifiers

1..545

/organism="Lycopersicon esculentum"

/cultivar="Rio Grande Flor"

/db\_xref="taxon:4081"

/clone="CLET3J21"

/clone\_lib="tomato mixed elicitor, BTI"

/tissue\_type="leaf"

/dev\_stage="4-6 week old plants"

/lab\_host="XLI-Blue MRF"

/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; CLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisocoumaric acid, BPH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

BASE COUNT 159 a 118 c 110 g 158 t  
ORIGIN

Query Match 22.8%; Score 204.2; DB 64; Length 545;  
Best Local Similarity 70.0%; Pred. No. 9.8e-40;  
Matches 275; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 313 tatagtgaactctacaaaacgagattcaaggaactatcctctgagcttggaaatctg 372  
DB 68 TATCTGAGCTTTACAAAATAATATTCAGGAAACCATCCCTAAGAGCTCGTAACTTG 127  
QY 373 aagagctctaatcagcttggatctgtacacaaacatctcacgcggaaatccatcttct 432  
DB 128 AAGAGCTTATTAAGTCTGTGATCTGTACAAACAAATATTTGGGACAAATCTCACTTCA 187  
QY 433 ttggaaatgtgaagcttcttcttgcggtctaaagaaacggaatgacgcgtctact 492  
DB 188 CTGGAAACCTGAAGAACTTTCTTTCGCTTAATGATTAACAACTAAGACGACCA 247  
QY 493 attctagaagaactcacagcttatctcaagccttaagatgtgtatgtctcaggaatgat 552  
DB 248 ATCCCAAGAGAACTTCTAGCATTTCTACCCGAAAGTTGGAGATGCTCAATTAATGAT 307  
QY 553 ttgtggaacaattccagtagaagacatttggacaatttggacaatttggacaatttgg 612  
DB 308 TTGTGGAACAATTCCTACTCTGTCTCATTTTGACATATT 367  
QY 613 aacaacctgagattggaggacagaactactaggcttgcgagctatgacacaattgc 672  
DB 368 CACAATCTCGAGTGAAGTCCGAGTGTGTTGGAGCTCGCTACTAGACACCAATGTC 427



```

Qy 673 acttaaaagaagttgaagaacctataagaag 705
Db 428 TCTTAGAAGCTTGATGAAATAGCTTCAGAGAAG 460

RESULT 15
LOCUS N65416 305 bp mRNA EST 05-JAN-1998
DEFINITION 20436 lambda-PRL2 Arabidopsis thaliana cDNA clone Z28M9T7, mRNA
sequence.
ACCESSION N65416
VERSION N65416.1 GI:1217042
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.
1 (bases 1 to 305)
Neman,T., deBrujin,F.J., Green,P., Keegstra,K., Kende,H.,
McIntosh,L., Onilogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
COMMENT On May 8, 1995 this sequence version replaced gi:801103.
Contact: Thomas Neman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ldm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1.305
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="Z28M9T7"
/note="Vector: lambda Z1p-lox; Site_1: Sal; Site_2: Not;
lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda Z1p-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA."
BASE COUNT 81 a 68 c 63 g 81 t 12 others
ORIGIN

Query Match 22.7%; Score 202.6; DB 25; Length 305;
Best Local Similarity 78.3%; Pred. No.2,3e-39;
Matches 235; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 382 atcagttgagctgtacacaacaaatctacacgggaaatcccatctcttttgggaaaa 441
Db 1 ATCAGCTTGCTGTGTCACAAACAATCTAACAGGGATAGTCCCACTTWTGSGAAAA 60

Qy 442 ttgaagtcactgttttttggcggttaacgaacacgattgacgcggtctatctcttctaga 501
Db 61 TTNAAGTTCGTCTTTTAACGCTTAATGACACCGATTGANGGGCCAAATCCTTAGA 120

Qy 502 gaactcacagttattcaagccttaaggttgatgtctcagggaatgatttgttggga 561
Db 241 AGTTGAGGGGCCCGGAATTACTGAGNTTGCAGAGNTTGGCAGCTTACTGACCTGAAAA 300

```

```

Db 121 GCATCTACTGCAATCCCAAGCNTTAAGTNGTTGATGCTCTCAGCAATGATTTGCTGGA 180
Qy 562 acaattccagtagaagagacctttgaacacatctcctatgcaaaacttggagaacacctg 621
Db 181 ACAATCCCAACAACGAGACTTTTGCTCAGATTCCTTTACAGAACTTNNGNACAAACCG 240

Qy 622 agatttgagggagaccacactaggtcttggagctatgacccaattgcacttaaaaa 681
Db 241 AGTTGAGGGGCCCGGAATTACTGAGNTTGCAGAGNTTGGCAGCTTACTGACCTGAAAA 300

```

Search completed: June 23, 2000, 19:06:42  
Job time: 27435 sec



•  
•  
•  
•



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2000, 00:23:23 ; Search time 12463.9 Seconds

(without alignments)  
-82.966 Million cell updates/sec

Title: US-09-180-798-30

Perfect score: 1063  
Sequence: 1 tcgaccaccgcgtccgacga.....tgcacaaaaaaaaaaaaa 1063

Scoring table: IDENTITY\_MTC  
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_bal.\*  
2: gb\_bal.\*  
3: gb\_bal.\*  
4: gb\_bal.\*  
5: gb\_bal.\*  
6: gb\_bal.\*  
7: gb\_bal.\*  
8: gb\_bal.\*  
9: gb\_bal.\*  
10: gb\_bal.\*  
11: gb\_bal.\*  
12: gb\_bal.\*  
13: gb\_bal.\*  
14: gb\_bal.\*  
15: gb\_bal.\*  
16: gb\_bal.\*  
17: gb\_bal.\*  
18: gb\_bal.\*  
19: gb\_bal.\*  
20: gb\_bal.\*  
21: gb\_bal.\*  
22: gb\_bal.\*  
23: gb\_bal.\*  
24: gb\_bal.\*  
25: gb\_bal.\*  
26: gb\_bal.\*  
27: gb\_bal.\*  
28: gb\_bal.\*  
29: gb\_bal.\*  
30: gb\_bal.\*  
31: gb\_bal.\*  
32: gb\_bal.\*  
33: gb\_bal.\*  
34: gb\_bal.\*  
35: gb\_bal.\*  
36: gb\_bal.\*  
37: gb\_bal.\*  
38: gb\_bal.\*  
39: gb\_bal.\*  
40: gb\_bal.\*  
41: gb\_bal.\*  
42: gb\_bal.\*  
43: gb\_bal.\*  
44: gb\_bal.\*

45: gb\_bal.\*  
46: gb\_bal.\*  
47: gb\_bal.\*  
48: gb\_bal.\*  
49: gb\_bal.\*  
50: gb\_bal.\*  
51: gb\_bal.\*  
52: gb\_bal.\*  
53: gb\_bal.\*  
54: gb\_bal.\*  
55: gb\_bal.\*  
56: gb\_bal.\*  
57: gb\_bal.\*  
58: gb\_bal.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1063	100.0	1063	5	A67825 Sequence 30
2	912.4	85.8	1106	5	A67817 Sequence 22
3	841.8	79.2	981	5	A67819 Sequence 24
4	769.6	72.4	789	5	A67821 Sequence 26
5	361	34.0	894	5	A67823 Sequence 28
6	211.8	19.9	2089	5	A67827 Sequence 32
7	155	14.6	936	8	SB062279 Sorghum bic
8	103.4	9.7	1755	8	DC093048 Daucus caro
9	103.4	9.7	1815	5	A67797 Sequence 2
10	99	9.3	4604	7	LEIRGENE X95269 L. esculentu
11	84.8	8.0	101284	8	ATAC009991 Arabidops
12	81.6	7.7	116944	8	AC008017 Arabidops
13	81.4	7.7	5033	8	IN077888 Ipomoea nil
14	79.8	7.5	15153	50	AC021198 Arabidops
15	79.2	7.5	134402	8	ATAC011765 Arabidops
16	78.8	7.4	100887	50	ATAC011620 Arabidops
17	78.8	7.4	108355	50	AC005957 Arabidops
18	78.6	7.4	133840	8	AC003981 Complete
19	76.8	7.2	115229	44	AC015446 Arabidops
20	76.4	7.2	23143	8	F15P11 AF160760 Arabidops
21	75.8	7.1	142418	7	AP000815 Oryza sat
22	75.4	7.1	3176	5	E12705 Arabidops
23	75.4	7.1	3176	8	ATU47029 Arabidops
24	75	7.1	5940	8	OL072727 Arabidops
25	74.6	7.0	84440	7	AB010075 Arabidops
26	74.6	7.0	118086	7	ATF6H11 Arabidops
27	72	6.8	87434	7	AB010698 Arabidops
28	71.4	6.7	3573	5	A57133 Arabidops
29	71.4	6.7	5368	8	U42445 Lycopersico
30	71.4	6.7	6471	5	A57130 Arabidops
31	71.4	6.7	6471	8	U42444 Lycopersico
32	71.4	6.7	61712	7	AB011061 Arabidops
33	70.8	6.7	4081	5	A67815 Arabidops
34	70.8	6.7	6695	5	A67796 Arabidops
35	70.6	6.7	98471	8	AC012654 Arabidops
36	70.6	6.6	2278	8	SBRUK1 Y14600 Sorghum bic
37	70.6	6.6	2786	7	AB029327 Nicotiana
38	70.6	6.6	3979	5	A67428 Arabidops
39	70.6	6.6	3979	5	A67429 Arabidops
40	70.6	6.6	3979	5	AF053993 Lycopersi
41	70.6	6.6	4123	8	A67434 Arabidops
42	70.6	6.6	4123	8	AF053998 Lycopersi
43	70.2	6.6	9424	8	OSU72724 Oryza sativ
44	70.2	6.6	52717	7	AB019227 Arabidops
45	70	6.6	101284	8	ATAC009991 Arabidops

## ALIGNMENTS



RESULT	1				
LOCUS	A67825	1063 bp	DNA		PAT
DEFINITION	Sequence 30 from Patent WO9743427.				05-MAY-1999
ACCESSION	A67825				
VERSION	A67825.1	GI:4756647			
KEYWORDS					
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1063)				
AUTHORS	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.				
TITLE	PRODUCTION OF APOMICETIC SEED				
JOURNAL	Patent: WO 9743427-A 20-NOV-1997;				
	CIBA GEIGY AG (CH)				
FEATURES	Location/Qualifiers				
source	1..1063				
	/organism="unidentified"				
	/db_xref="taxon:32644"				
	106..762				
CDS	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAB42252.1"				
	/db_xref="gi:4756648"				
	/translation="MASRNYKWELEFAASLITLTALIHVEANSEGDAIYALRSLTDP				
	DHYVLSMDPLVNPCTWFHVTCNODNRYTRVDLGNSSNGSLAPEGKLEHLYELEY				
	KNNQGTIPSELGKLNKILSDLENNNLGTGVPNSLKIKSLIVFLRLNDNRLTGPDR				
	ALTAIRPRLKVVYDSSNDLCGTPPTNGFPAHITPLDNFNNRRLBEGPELLGLASTYDNTCT				
BASE COUNT	313 a 242 c 206 g 302 t				
ORIGIN					

Query Match	100.0%	Score 1063	DB 5	Length 1063
Best Local Similarity	100.0%	Pred. No. 2.3e-263		
Matches 1063	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	1	tcgaccacacgcgtccgcagaaacccctaatttgccttcctcatccttgcttcgaagaattact	60	
Db	1	TCGACCAAGCGGTCCGAGAAACCTTAATTTTGGCTTCGTCATCTGTTACAGAAATTACT	60	
OY	61	caaatctctattagattactctctcttcgcacctcgatagtcacatgagcgtctcgaac	120	
Db	61	CAAAATCTTATAGATTACTCTCTTCGACCTCCGATAGCTACATATGCGTCCTCGAAGC	120	
OY	121	tatgggtggagactcttcgcagctcttcagctcgttaaccacacttgcttgatttaacctg	180	
Db	121	TATGGGTGGAGACTCTTCGACCTTCGATTAACCTTAAGCTTGTATTACCTGGGC	180	
OY	181	gaagcaaacctccgaagagatgctctttagcctcttcgcggaggtttaaagatccgac	240	
Db	181	GAAAGCAACTCCGAAAGAGATGCTCTTACGCTCTTCGGCGAGTTAACATATCCGGAC	240	
OY	241	catgtcttcagagctggagatcccaactctgttaactcttgtaacctggtgtccatgcc	300	
Db	241	CATGTCTTCCAGAGCTGGGATCCAATCTGTTTAATCTTTGACCTGGTTCATGTCACCC	300	
OY	301	tctaaccaagaaacacgcgtcaactcgttgtgtgttggtgggaattcaaacctccttgaa	360	
Db	301	TCTTAACCAAGAACACCGGCTCACTCGTGTGGATTGGGGAATTCAAACCTCTCTGGACAT	360	
OY	361	cttgagcctgagcttggagagcttgaacattcacagtaactagagctcaaaaaaac	420	
Db	361	CTTGCGCCTGAGACTTGGAGAGCTTGAACTTTACAGTATCTAGAGCTCTACAAAAAAC	420	
OY	421	atccaaggaactataccttcgcgaacttggaaatctgaaagatctcaacacacac	480	
Db	421	ATCCAAGGAACATATACCTTCCGAACCTTGGAAATCTGAAGATCTCATCAGCTTGATCTG	480	
OY	481	tacaacaacaactcttcacaggaatagtctccaactcttgggaaattgaagctctcgtgc	540	
Db	481	TCAACAACAACCTTACAGGAATAGTCTCCACTCTCTTGGGAAAAATTGAAGTCTCGGTC	540	

Query Match	85.8%	Score 912.4	DB 5	Length 1106
Query Local Similarity	95.0%	Pred. NO. 1.3e-224		

  

BASE COUNT	331 a	258 c	206 g	311 t
ORIGIN				

  

RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	CDS
A67817	A67817	Sequence 22 from Patent WO9743427.	A67817	A67817.1	GI:4756639	unidentified.	unclassified.	1 (bases 1 to 1106)	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.	PRODUCTION OF APOMITIC SEED	Patent: WO 9743427-A 20-NOV-1997;	Location/Qualifiers	
												1..1106	
												/organism="unidentified"	
												/db_xref="taxon:32644"	
												142..798	
												/note="unnamed protein product"	
												/codon_start=1	
												/protein_id="CAB42244.1"	
												/db_xref="GI:4756640"	
												/translation="MASRYKWELEFASLTLTALILHVEANSSEDAIYALRSITDIDT	
												DIVLDSMPETLVNPTCMFHWYGNODNVRVVDGNSNLSGHLAPKLEHLOYLEL	
												KNNIGGTIPSELGNKLNKLNISDLYNNNTLTIVPTPIFGKLSYFPLRLNDRLTGP	
												HSRQSAKRVVDVSSNDLCGTIPNGFAHPIPLQNFENNPRLGEPBLGLASTYD	







Db 537 TCTTTTACGGCTTAATGACAACCGATGACCGGTCAATCCCTAGACACTACAGGCAA 596  
Qy 559 tcccaagccttaagttgttgcctcaagcaatgatttctgtggaacattcccaaaa 658  
Db 597 TCCCAAGCCTTAAAGTTGTGACGCTCAGCAAGATTTGTGGAAACAATCCCAACAA 656  
Qy 659 acgagccttctgcacattcccttacagaaccttgagaacacccgggttggaggagac 718  
Db 657 ACGGACCTTGTCTACACTCTCTTACAGAACTTGGAAACACCGGAGATTGGAGGAC 716  
Qy 719 cggaaattacccgtctcgaagctacgacactgaactgacactggaataattggcaaac 778  
Db 717 CGGAATTAAGTCTGCTTCAAGCTACGACACTGAACCTGCACTGGAACAAACGAAAAAC 776  
Qy 779 tgaataagaagattgggggtgacctgtgaagaaccttacaccttatacaattatca 838  
Db 777 TGAATAAGAAATGGGGGGGAGACCTTGTAGAAACACTTCAACCTTAATCAATATCA 836  
Qy 839 cactcactatgtaaatatataatgtagtccaa--aaaaaaatgaagatcgaatc 896  
Db 837 CATCTATATATTAATTAAGTATATATATAGTAAACAAAAAATGAATCGAATC 896  
Qy 897 agtaataatcctgtctcaattggaaccttgaggctgtgtatgtaaatcttctaact 956  
Db 897 GGTAAATATCATCTGCTCTCAATTGAGAACTCGAGGTC--TGTATGTAATAATTCTAAT 954  
Qy 957 gggacttcggctactgtaact 977  
Db 955 GCGATTTTCGCTTAATTAAT 975

RESULT 4  
LOCUS A67821 789 bp DNA PAT 05-MAY-1999  
DEFINITION Sequence 26 from Patent WO9743427.  
ACCESSION A67821 GI:4756643  
VERSION A67821.1  
KEYWORDS  
SOURCE .  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 789)  
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE PRODUCTION OF APOMITIC SEED  
JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)

FEATURES  
source Location/Qualifiers

CDS  
1..789  
/organism="unidentified"  
/db\_xref="taxon:32644"  
<2..664  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42248.1"  
/translation="RPTRPYRWELFPAALILITLALIHVEANSSEDAVYALRSIT  
DPDYLQSDPFLVPCWTFVTCNODNRVRYVDLGSNSGHLAPELGKLEHLYE  
LYKNNGIETISELGNLKNLISLDLYNNNGIYPTSLGKSLVFLRNDNRITGPI  
PRALTAISLVVYSSNDLGTIPTNCPFAHIFLQNFENNRPLEGPELGLASIDYN  
CT"

BASE COUNT 235 a 191 c 156 g 207 t  
ORIGIN

Query Match 72.4%; Score 769.6; DB 5; Length 789;  
Best Local Similarity 98.9%; Fred. No. 7.6e-186;  
Matches 775; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 103 cacatgagctcgaacatcggtgagctcttcgacgctgttaacttaacctta 162  
Db 5 CCCAGCGCTCCGGAACATCGGTGGAGCTCTTCGCGACTTCCTTAATCCTTA 64  
Qy 163 gcttgatcactggtcgaagcaactccgaaggagatgctctttagcgtcttcg 222

Db 65 GCTTTATTCACCTGGTTCAGAGAAACTCCGAAGGAAATGCTTTAGCGCTTGCCGG 124  
Qy 223 agttcaagatcggagacatggttccagagctggatccaactctgttaactctgt 282  
Db 125 AGTTTAAAGATCCGAGCATTGTTCCAGAGCTGGGATCCAACTTGTTAATCTTGT 184  
Qy 283 acctggtccatgtaaccgttaaccagacacgcgtcaactcgtgtgatttggagat 342  
Db 185 ACCTGGTCCATGTCACCTGTAAACAGAACACCGGTCACTCGTGTGATTTGGGAAAT 244  
Qy 343 tcaaacctctcggacatcttcgcctgagcttgggaagcttgaacattacaagatca 402  
Db 245 TCAACCTCTCTGACATCTTGCGCCTAGCTGGGAAGCTTGAACATTATACATCTA 304  
Qy 403 gagcttcaaaaaaacatccaaggaactatacttcggaacttggaaacttgaagat 462  
Db 305 GAGCTTACAAAAAACATCCAGAGAACTATACCTTCGAACCTTGGAAATCTGAAGAT 364  
Qy 463 ctcatgactggtatctgtacaacaacatcttaagagatagttccactcttggga 522  
Db 365 CTGATGAGCTTGATCTGTACAAACAAATCTTACAGGATAGTCCACTTCTTGGA 424  
Qy 523 aaattgaagctcgtgctcttcttaacggttaatgaacacggaattgacggggcaatcc 582  
Db 425 AAATGGAAGTCTGCTCTTTTACGGCTTAATGACAAACGATGACGGGGCAATCCCT 484  
Qy 583 agagcaactcaatcccaagccttaagttgtgtatgctcaagcaatgatttgt 642  
Db 485 AGAGCACTCACTCAATCCCAACCCCTTAAGTGTGAGTGTCAAGCAANTATTGTGT 544  
Qy 643 ggaacaatcccaacaacgaccttctgctacatctcttaacgaacttgaagaaac 702  
Db 545 GGAACATCCCAACAACGAGACCTTTGCTCAGATTCCTTACAGACTTTAGAAAC 604  
Qy 703 ccgaagttggaggagccgaatctgctcgtcaagctgaagcaactgaactcgtga 762  
Db 605 CCGAGGTGGAGGAGCCGGAATTAATGCTGCTGCAAGCTAGACACTGACACCTGA 664  
Qy 763 aaaaattggcaaaactggaatgaagaattgggggtgtaacttgaagaactcacc 822  
Db 665 AAAAAATTGGCAAAACCTGAAAGAAATGGGGGTACCTTGTAAACACCTTACC 724  
Qy 823 accttaacaatacactactactatgtaataagatataatgtaagtcacaaaaaaa 882  
Db 725 ACTTTTCAAAATCACTACTACTATGTAATATATATATATGTAAGTCAAAAAA 784  
Qy 883 tga 886  
Db 785 AAAA 788

RESULT 5  
LOCUS A67823 894 bp DNA PAT 05-MAY-1999  
DEFINITION Sequence 28 from Patent WO9743427.  
ACCESSION A67823  
VERSION A67823.1 GI:4756645  
KEYWORDS .  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 894)  
AUTHORS De,V.S., Schmidt,E.D., Van,H.G. and Hecht,V.F.  
TITLE PRODUCTION OF APOMITIC SEED  
JOURNAL Patent: WO 9743427-A 20-NOV-1997;  
CIBA GEIGY AG (CH)  
FEATURES  
source Location/Qualifiers  
1..894  
/organism="unidentified"  
/db\_xref="taxon:32644"  
<1..678  
/note="unnamed protein product"



[illegible]

ACCESSION	A67827
VERSION	A67827.1
KEYWORDS	GI:4756649
SOURCE	' thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS	Euhariyoti; Viridiplantae; Streptophyta; Magnoliophyta; eudicotyledons; core
JOURNAL	eudicots; rosidae; eurosids II; Brassicales; Brassicaceae;
FEATURES	Arabidopsis.
Source	1 (bases 1 to 2089)
CDS	De,V.S., Schmidt,E.D., Van,H.G. and Hecht,Y.F. PRODUCTION OF APOMICRITIC SEED Patent: WO 9743427-A 20-NOV-1997; CIBA GEIGY AG (CH) Location/Qualifiers 1..2089 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="SERK GENE CDNA" 195..2072 /note="unnamed protein product" /codon_start=1 /protein_id="CAB42254.1" /db_xref="GI:4756650" Translation: "MESSYHYVEILLSTLLPNHSLMTASNLAEDELALTRLTVLDNPN NVLQSDPLTNPCTWTFHWNNNSYIRDLGAELSGHLYPELGLKNLOYLEYS NTNGIPSPNNGTNTNVSLDIYNRSQGPSEIQLSKTRFLRNUNNSISGTPMS LNITTLQYLVDLSNNRLSGSYVPDNGFSLTLPISFANNLDGCVISHPCGSPFPSS PPPTLPPPVATPSGVIGTGAIAAGVANAGAALPRAAIAFAWMRRSPIDIFDVPA EEDDEVHLGOLKRFSLRELVOVASDFENKNIILGGFGKYKGKLADGTVAVKRE ERTGEJGELOTEVEEMTSMAVRMLRLRGCMPTERLIYYPNANGSVASCLEBP PSCPDPMPETPKRIALSGARLSYLSHPDCPI.IHRDYKANIIILDEFEAVVDGFI AKLMDDYHTVTAVRTIGHIAPESYTGSSSKTPOFGIMLLELILGOAFDIA PLADDVVMIIDMWIKLEKKLEPMIVDPLQCTYERRELDQVQVALICTOSPMRR PKMEVRYMLEBDELAKRWDMQVVELIREETIDSPNDSMILDSTYINLAVELSGP R"
BASE COUNT	569 a 419 c 502 g 599 t
ORIGIN	
Query Match	19.9%; Score 211.8; DB 5; Length 2089;
Best Local Similarity	63.0%; Pred. No.3e+4;
Matches 327; Conservative	0; Mismatches 192; Indels 0; Gaps 0;
Ox	184 gcaaacctccgaaggagatcgtcttctaagctcttcgcccgaagttaacagatccggaccat 243
Dd	270 GCTAACTTTGGAAAGCGATGACTTGGATACTTTGAGGGATTACTCTAGTCATCAACAAT 329
Oy	244 gtctccagaagcttgatccaactcttvttaaccttbtacctggttccatgtaacctgt 303
Dd	330 GCTTCGCAGAGCTGGCAATCTACGCTAGTAGAATCTTCCACATGCTTCACATGCTCACTTGC 389
Oy	304 aaacaagaacaacgcggtctacgtcgtytgtatttgggaaattcaaacctctccggaaattt 363
Dd	390 AACCAACGAGAACAGTGTATTAAGAGTTGATTTGGGGAATGCAAGATTATCTGECATTTA 449
Oy	364 gcgccttgagcttgggaaagcttgaaatttaacagta.ctagaagctctacaacaaacaacatc 423
Dd	450 GTTCCAGAGCTTGTGGTGTCTCAAGAAATTTGCGATTTGGAGCTTTACAGTAACAACATA 509
Oy	424 caaggaactatacccttcogaactttgaaattcgaagaatcatcaagcttggattctgtac 483
Dd	510 ACTGCCCGCATTCCTAGTATCTTGGAAATCGAACAACCTTAGTAGTGGATTTGATTTTAC 569
Oy	484 aacaacaactcttacaggagtagtccccactctcttgggaaaaattgaagctctctgtcttt 543
Dd	570 TTAACACAGCTTCTCCGGTCTATCTCCGGAATCATTTGGGAAAAGCTTTCAAAGCTGAGATT 629
Oy	544 tttaggctttaagaacaccgatggcggggccaatccctcatagaagcaactcatgcaatcccca 603
Dd	630 CTCGGGCTTAAACAACAAGTCTCACTAGGCTCAATTTCTTCAATGCTACCAACATTTTACT 689
Oy	604 acccttaaatgctttcatgctctcaagcaataatgttctltgaaacaatccgaacaaagga 663



```

Db 690 ACCCTCAGGTGTAGATCATCAATACAGACCTCTGGGTGACTGCTTCCATCAATGCG 749
QY 664 cctttgtcacattcctttacagaacttggagaacac 702
Db 750 TCCCTCTCACTCTTCACACCCATCAAGTTTGGCTAATAAC 788

RESULT 7
SB062279 936 bp mRNA PLN 09-DEC-1996
LOCUS Sorghum bicolor leucine-rich repeat-containing extracellular
DEFINITION glycoprotein mRNA, complete cds.
ACCESSION U62279.1 GI:1710123
VERSION 062279.1
KEYWORDS
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum.
REFERENCE 1 (bases 1 to 936)
AUTHORS Hipskind,J.D., Nicholson,R.L. and Goldsbrough,P.B.
TITLE Isolation of a cDNA encoding a novel leucine-rich repeat motif from
Sorghum bicolor inoculated with fungi
JOURNAL Mol. Plant Microbe Interact. 9 (9), 819-825 (1996)
MEDLINE 9714217
REMARK Erratum: [[published erratum appears in Mol Plant Microbe Interact
1997 Mar;10(2):302]]
REFERENCE 2 (bases 1 to 936)
AUTHORS Hipskind,J.D.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-1996) John D. Hipskind, Botany and Plant
Pathology, Purdue University, West Lafayette, IN 47907, USA
FEATURES
source
1..936
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone="SLR"
12..755
/function="protein-protein ligand binding"
/note="leucine-rich repeat-containing extracellular
glycoprotein: contains six N-glycosylation sites
[IK(S/T)]"
/codon_start=1
/protein_id="AAC49559.1"
/db_xref="GI:1710124"
/translation="MAPOFAAGFLTGALATPASCNTGSDIYKORLAWEDPNVY
OSMNSTLANPCTWFHVTQNNNNFVIRVDLGNAGISGPLPLDAEIONQYELKNGI
NSIPETIGNLISLIDMLDLTGEITPTGTSYSTRYLKXONNLGTIPSPFGN
LTSLESTKQENSLSGATPASIGNIKALOFSLNDMLTGTVPKSPFLSFGNLTEL
NTDRNNLDGTRTSSGLRYTAIIQDALKTA"
36..77
/evidence="not_experimental"
BASE COUNT 247 a 232 c 207 g 250 t
ORIGIN

Query Match 14.6%; Score 155; DB 8; Length 936;
Best Local Similarity 58.2%; Pred. No. 1.2e-29;
Matches 272; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 187 aactcgaagaagatgctcttaccgctcttcgcgcgaagttaacagatcgcgaacattgt 246
Db 81 AACACTGAAAGTGACATACGTACAAAGCAAGGTTGGCATGGAGAGCAACCAACGATG 140
QY 247 ctccagagctggagatccaactctgttaactcttgcactgttccatgttccactgaac 306
Db 141 CTGACAGAGCTGGAATTCGACGCTTGCAATCCCTGCACCTGCTTCATGTCACCTGCAC 200
QY 307 caagacaaccgcgtcactcgtgtgatttggggaaattcaaacctctctggaactcttgcg 366
Db 201 AACAAATTAACCTTTGTCAATCCGCTGGATTTGGCAATGACAGCATCTCCGGTCTCTGTT 260

```

```

QY 367 cctgagcttggagaagcttgaacattacagtagtactagaagctctacaaaacaacatccaa 426
Db 261 CCAATATCTTCAGAAATTCAGAACTCCAGTACATGAGTGTATGGCAAGCTTGAAAC 320
QY 427 ggaactatattccgaacttgaatactgaataactcaactcagcttgatctgtacaac 486
Db 321 GGTTCGATTCAGAAACACAGGCAACCTGACATCTCATGATGATCTGATCTCTGGAC 380
QY 487 aacaacttaccaggaagtagtccactcttctgggaataatgaagctctgtgctttta 546
Db 381 AACCTTCTTACCGCGCAAACTCCAACTACGCTGTCTGTCTGACGACGCTGCATCTG 440
QY 547 cggcttatagcaacgattgacggggccaatccctagagcactcactcgaatcccaagc 606
Db 441 AGGTTGTACCAAAACACCTGACAGGCGCTATACATCATCTTTGGCAACCTGACTAGC 500
QY 607 cttaagttgtgtatgtctcaagcaatgtattgtgtggaacaatccc 653
Db 501 CTTCTGGAATCGAAGCTTCAAGAGATTCCTTGAAGCGCGCTATTC 547

RESULT 8
DCU93048 1755 bp mRNA PLN 28-JUN-1997
LOCUS Daucus carota somatic embryogenesis receptor-like kinase mRNA,
DEFINITION complete cds.
ACCESSION U93048
VERSION U93048.1 GI:2224910
KEYWORDS
SOURCE .
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; eucolecyledons; core
eudicots; Asteridae; eusterids II; Apiales; Apiaceae; Daucus.
REFERENCE 1 (bases 1 to 1755)
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A. and de Vries,S.C.
TITLE A leucine-rich repeat containing receptor-like kinase marks somatic
plant cells competent to form embryos
JOURNAL Development 124 (10), 2049-2062 (1997)
MEDLINE 97313247
REFERENCE 2 (bases 1 to 1755)
AUTHORS Schmidt,E.D., Guzzo,F., Toonen,M.A.J. and de Vries,S.C.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-1997) Molecular Biology, Agricultural University
of Wageningen, Dreyenlaan 3, Wageningen 6703 HA, The Netherlands
FEATURES
source
1..1755
/organism="Daucus carota"
/db_xref="taxon:4039"
34..1695
/note="SERK"
/codon_start=1
/product="somatic embryogenesis receptor-like kinase"
/protein_id="AAB61708.1"
/db_xref="GI:2224911"
/translation="MNRNSINILNTWQFDAYIDKYGVIMTELYSNNISGPIPSDIG
NLINIVSLDIYANSGPILPDLIGKILTRIRLRNNNSISGPIPSLIRITITLQYIDL
SNRLSGVPVNDGNSFSLTPIFANNLNCGVYTGPCGSPSPSPPTPSTVOP
PGONGPTGAIAAGVAAGAAALFAAPAMAMARRRRARHREFDVAEDPEVHALQK
RFSRLQVATDTFTSTILGRGGFKYKORLADGSIYAVKRLKEERTGGELQFQTEV
EMISMAVHRLRLRGCFMPTERLLVYVMANGSVASCLEROSPSEPLDKPTTKRI
ALGSARGLSYIHQHPKTIHRDYVARNILDEEPAVYDGLRIMDYKPTHTTA
VRGTGLYIAPETISTKSSKSEKIDVRYGIMLELITLGGAPLALANDDDVMDLVY
KSLKKEKLEMLVDPLENNYIDTEYEQILQVALCTQDSPEPERKMSVYVMLBGDG
LAERKMDQKVEVTHQDVELAPHRSEWILSDTNLHAFELSGPR"
BASE COUNT 506 a 347 c 407 g 495 t
ORIGIN

Query Match 9.7%; Score 103.4; DB 8; Length 1755;
Best Local Similarity 55.4%; Pred. No. 2.4e-16;
Matches 200; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

```



QY	363	tgagccttgagcttgggaagcttgaacatttacagtatcttagagctctcaaaaacaacat	422
Db	78	TGATGCTTTACCTTGACCAATATGGGGTTCCTTAATGACATTGGAGCTTTACAGCAATTAACAT	137
QY	423	ccaaaggaacataccttccagacttggaaactctgaagaaactcctcaactcattgacttctgta	482
Db	138	AAGTGGACCAATTCCTAGTGAATCTTGGGAATCTGACAAATTTGGAGCTTGGACCTATA	197
QY	483	caacaacaacttacaaggatagttcccaactctcttgggaaattgaagctctggtcctt	542
Db	198	CATGAATFAGCTTCTGTGGACCTATACCGGACACATTAGGAAAGCTTACAAAGCCTAAGATT	257
QY	543	tttagagcttaatgcaacogattgaaggggccaatccctaaagacactcactgaatccc	602
Db	258	CTTGCGCTCACAACACAGAGCTCTCTGGTCCCATTTCCAAATGCTACTAATATTAC	317
QY	603	aagccttaaaagttgtgtagtctcaagcaatgatcttgtggnacatcccaacaacog	662
Db	318	AAGCTTTCAAGTCGAGATTTATCAACAAATCGGCTATCAGGACCAAGTACCGGATTAAG	377
QY	663	accctttgctcacatctcctttaacagaactttgagacaacccgaggttggaggaacgga	722
Db	378	CTCATATTTCTTTGTTACACCTATACGTTTTGGCAATAATTGATTTATGTGGACCGGT	437
QY	723	a 723	
Db	438	A 438	

[illegible]

BASE COUNT	ORIGIN	FEATURES	LOCUS	RESULT
531 a	354 c	415 g	515 t	9
				67797
				DEFINITION
				Accession
				Version
				Keywords
				Source
				Organism
				carrot.
				Daucus carota
				Eukaryota; Vitidiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids II; Apiales; Apiales; Daucus.
				1 (bases 1 to 1815)
				De V.S., Schmidt, E.D., Van, H. G. and Hecht, V.F.
				PRODUCTION OF APOMICRIC SEED
				Patent: WO 9743427-A 20-NOV-1997;
				CIBA GEIGY AG (CH)
				Location/Qualifiers
				1..1815
				/organism="Daucus carota"
				/db_xref="taxon:4039"
				94..1755
				/note="unnamed protein product"
				/codon_start=1
				/protein_id="CAB42241.1"
				/db_xref="gi:4756624"
				/translation="MNRSNINILNYMOETDAYLDKYGVLMTLELYSNNSISGPISPDLDGLNITNYSILDIYMNRSQGIPTDGLKRLRFLRFLMNNNSISGPIPMSTNITITLOVIDISNRRISGPIPDGSGFSLSLTPISFANNINILGIPYGRPCRGSPSPFPPTPIPTSVQPGNGETGAIAAGVAAAGALFLPAAPAMAFVRRRRKRFEPVPAEDPEVHLSQLELRSLRLQVADITFSTILLRGSGFGRVYKRWAGSGLVAYKRLKEERTPGEGLOFQTEVEMSAVYRNRLRLRGFCMPTPLRLVLYPYMANGSVASCILREBPSEPLDPTPKRRLALGASGLSYLHDHCDPKIRRDVYKANIILDEEFVAVGDFGLARMDYKOTHTVTRVSTLGIYIPEYLSKSEKTDVGYGKILDEELITGORAFLARLANNDDVNLIDMDKSLIKERKLEMTVPDPLFNNTIDPEVGLIIOVALICGQSGPMRPRPMSYVMTLEGDDILAKRNDENQVEYIHQDELAPHRSTSEMILSDTNDLHAFELSGPR"

Query Match	9.7%;	Score 103.4;	DB 5;	Length 1815;
Best Local Similarity	55.4%;	Pred. No. 2.4e-16;		
Matches 200;	Conservative	0;	Mismatches 161;	Indels 0;
				Gaps 0;

LOCUS	LEIRGENE	10	RESULT
DEFINITION	L.esculentum LRP gene.	PLN	10-OCT-1996
ACCESSION	X95269		
VERSION	X95269.1 gi:1619299		
KEYWORDS	LRP gene; LRP protein.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; Solananae; Solanales; Solanaceae; Solanum; Potatoes; Lycopersicon.		
AUTHORS	1 (bases 1 to 4604)		
TITLE	Vera, P.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (18-JAN-1996) P. Vera, Universidad Politecnica de Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino de Vera 14, E- 46022 Valencia, SPAIN		
AUTHORS	2 (bases 1 to 4604)		
TITLE	Tornero, P., Mayda, E., Gomez, M. D., Canas, J., Conejero, V. and Vera, P.		
JOURNAL	Characterization of LRP, a leucine-rich repeat (LRP) protein from tomato plants that is processed during pathogenesis		
MEDLINE	Plant J. 10 (2), 315-330 (1996)		
FEATURES	96367673		
SOURCE	Location/Qualifiers		
exon	1..4604		
gene	/organism="Lycopersicon esculentum"		
	/cultivar="VFN8"		
	/db_xref="taxon:4081"		
	/tissue_type="leaf"		
	<1..345		
	/number=1		
	113..4355		
	/gene="LRP"		
CDS	join(113..345,2381..2452,2565..2708,2928..2999,3081..3152,4283..4355)		
	/gene="LRP"		

```

protein_cdn=Q96456.1
protein_id=Q96456.1
db_xref=GI:1619300
/db_xref=SPRKBMB:Q96477

```



```

/translating="MEAFKFTOSLFIKIMLAVLAVAVAVKNSGEGALYALRSL
SDPONTQSDPULVYPCPTFFHTCNGNDQVYTVYDGNKSLSHLYPELGEKHLOYL
ELYKNNTQGTIPKELGNLSLISLDLYNNISSTITPSGNLNVFLKLNDRKLTGP
IPRELNTSISLTKVDVSNNDLCGTIPTSGFPHIPLNNEHNPRLGPELGLATYDI
NCS"
Intron      346..2380
             /gene="Lrp"
             /number=1
             /number=1
exon        2381..2452
             /gene="Lrp"
             /number=2
             /number=2
Intron      2453..2564
             /gene="Lrp"
             /number=2
             /number=2
exon        2565..2708
             /gene="Lrp"
             /number=3
             /number=3
Intron      2709..2927
             /gene="Lrp"
             /number=3
             /number=3
exon        2928..2999
             /gene="Lrp"
             /number=4
             /number=4
Intron      3000..3080
             /gene="Lrp"
             /number=4
             /number=4
exon        3081..3152
             /gene="Lrp"
             /number=5
             /number=5
Intron      3153..4282
             /gene="Lrp"
             /number=5
             /number=5
exon        4283..>4604
             /number=6
             /number=6

BASE COUNT  1311 a      758 c      888 g     1647 t
ORIGIN
Query Match      9.38; Score 99; DB 7; Length 4604;
Best Local Similarity 77.48; Pred. No. 3.3e-15;
Matches 120; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Oy 178 gtcgaagcaactccgagagagatctcttcaagcttcgcgcgaggttaacagatccg 237
    || || || || || || || || || || || || || || || || || || || || ||
Db 194 GTAAGGGGGAATTCAGAGGGGATCTGTGACGCCCTCCCGAGATTATCGACCCG 253

Oy 238 gaccatgtctccagagctgggatacactctgtttaacctgtacctggttccatgtc 297
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 254 GGTAAAGTGTTCAGAGAGCTGGGATCCAAATCTGTGTACCTGTGACCTGTTCAATGTC 313

Oy 298 accttaaccagaacacgcgcgtccactcgtgtgga 332
    || || || || || || || || || || || || || || || || || || || || ||
Db 314 ACTTCAACGGAGATATCAAGTACTCTGCTGTGTA 348

RESULT 11
LOCUS      ATAC009991.101284 bp      DNA      PLN      08-Oct-1999
DEFINITION Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence,
complete sequence.
ACCESSION  AC009991
VERSION    AC009991.3
KEYWORDS   HTG.
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons; core
            eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
            Arabidopsis.
REFERENCE  1 (bases 1 to 101284)
AUTHORS   Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B.,
            Rominig,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
            Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.

```

```

TITLE      Arabidopsis thaliana chromosome III BAC F9f8 genomic sequence
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 101284)
AUTHORS    Lin,X. and Kaul,S.
TITLE      Direct Submission
JOURNAL    Submitted (09-SEP-1999) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
            3 (bases 1 to 101284)
AUTHORS    Lin,X.
TITLE      Direct Submission
JOURNAL    Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
            On Oct 8, 1999 this sequence version replaced gi:5902413.
COMMENT    Address all correspondence to:
            Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr.
            Rockville, MD 20850, USA
            e-mail: xlin@tigr.org
            BAC clone F9f8 is from Arabidopsis chromosome III and is near the
            molecular marker g4547.
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.
            Genes were identified by a combination of three methods: Gene
            prediction programs including GRAIL (available by anonymous ftp
            from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of
            Washington), Genscan (Chris Burge,
            http://genome.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
            (http://www.cbs.dtu.dk/netpene/obsnetpene.html), searches of the
            complete sequence against a peptide database and the Arabidopsis
            EST database at TIGR (http://www.tigr.org/tdb/at/at.html).
            Annotated genes are named to indicate the level of evidence for
            their annotation. Genes with similarity to other proteins are named
            after the database hits. Genes without significant peptide
            similarity but with EST similarity are named as 'unknown' proteins.
            Genes without protein or EST similarity, that are predicted by more
            than two gene prediction programs over most of their length are
            annotated as 'hypothetical' proteins. Genes encoding tRNAs are
            predicted by tRNAScan-SF (Sean Rddy,
            http://genome.wustl.edu/eddy/tRNAScan-SF/). Simple repeats are
            identified by RepeatMasker (Arjan Smil,
            http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
            genomic sequence that are not annotated as genes but have predicted
            exons by GRAIL are annotated as misc features.

FEATURES             Location/Qualifiers
     source            1..101284
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="III"
                     /map="g4547"
                     /clone="F9f8"
     misc_feature      77..108
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature      124..259
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature      complement(347..408)
                     /note="exon predicted by xgrail, quality
                     excellent_shadowexon"
     misc_feature      355..594
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature      672..752
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature      854..973
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature      complement(1182..1357)
                     /note="exon predicted by xgrail, quality
                     marginal_shadowexon"
     misc_feature      complement(1516..1694)
                     /note="exon predicted by xgrail, quality
                     excellent_shadowexon"
     misc_feature      1546..1666
                     /note="exon predicted by xgrail, quality good"

```



misc_feature	1747..1838	/note="exon predicted by xgrail, quality excellent"
misc_feature	/note="exon predicted by xgrail, quality good_shadowexon"	
misc_feature	2114..2246	/note="exon predicted by xgrail, quality excellent"
repeat_region	complement(2975..3024)	
mRNA	repeat_family="(CAA)n"	
gene	join(3070)..3325,3565..3642,4094..4234,4338..4414,4506..4655,4800..>5013)	
gene	/gene="F9F8.1"	
gene	3070..>5013	
gene	/note="F9F8.1"	
gene	/note="similar to unknown protein GB:AAC27167 [Arabidopsis thaliana]"	
CDS	join(3215)..3325,3565..3642,4094..4234,4338..4414,4506..4655,4800..5013)	
gene	/gene="F9F8.1"	
gene	/note="unknown protein"	
gene	/codon_start=1	
gene	/protein_id="AAFP01505.1"	
gene	/db_xref="GI:6016678"	
gene	/translation="WGPAPQIVLVFGSSIVYVFGSGHGWGALISEYARKADIIINGVYGNSSRLAEVDQVPEPRDAVQSLIYVFGSGDSKAPSSGLCPHVLTEVDMMKIALHLSIDSFRIIFLSPVDYAEARONOSPYLEVTRINDLCKTSDACVEICDLGLAEVLDLFTFEOKRADMKVCTDGIHLSDGSKITVAEILIRVYKEMEMHSLHM KMPLEFADDSYDLYSADGKOTVNSSEMTYFWEQMD"	
repeat_region	complement(5513..5597)	
repeat_region	repeat_family="(GA)n"	
misc_feature	complement(5527..5579)	
misc_feature	/note="exon predicted by xgrail, quality excellent_shadowexon"	
mRNA	join(5589)..5795,5897..5929,6013..6235,7264..7387,7467..7932)	
gene	/gene="F9F8.2"	
gene	5589..7932	
gene	/note="F9F8.2"	
gene	/note="similar to nucleic acid binding protein Alf1n-1 GB:AAA20093 [Medicago sativa]"	
CDS	join(5656)..5795,5897..5929,6013..6235,7264..7387,7467..7727)	
gene	/gene="F9F8.2"	
gene	/codon_start=1	
gene	/product="putative nucleic acid binding protein"	
gene	/protein_id="AAFP01506.1"	
gene	/db_xref="GI:6016679"	
gene	/translation="MAAAVSSNPRTVEIFRKYDSARRAALLRALTKDVDFYSCQCP EKENCLVGHPESEENKRLPAEVEPTELPEBALGYNPARDMQKMDLSIVAHSDCW LLSVYFIRGARLRNNEKRRKRLSLINDLPTLFDVYIRAKMDKNPSSDSGSKHGKTK NSIDQGTSTPKLMESETEEEDEEDGDTLCSGGHYNEEFWICDVCERWYHGK KCVKITPAAESIKYKCPCCAKKRG"	
repeat_region	complement(7711..7760)	
repeat_region	repeat_family="(CAR)n"	
repeat_region	complement(7810..7847)	
repeat_region	repeat_family="(GAAA)n"	
mRNA	8515..8572	
gene	/repeat_family="(PAA)n"	
gene	complement(join(<9426..9674,10198..10528,11098..11491..>11865))	
gene	/gene="F9F8.3"	
gene	complement(<9426..9674,10198..10528,11098..11491..>11865)	
gene	/gene="F9F8.3"	
gene	/note="similar to leucoanthocyanidin dioxygenase GB:BA20143 [Perilla frutescens]"	
CDS	complement(join(9426..9674,10198..10528,11098..11491..11865))	
gene	/gene="F9F8.3"	
gene	/codon_start=1	
gene	/product="putative leucoanthocyanidin dioxygenase"	
gene	/protein_id="AAFP01507.1"	
gene	/db_xref="GI:6016680"	
gene	/translation="WNNDELKIESKTLCLNDQDEVRKIDNNHMSDQDKRIEINKSG LGEKPELIVRQSLAESNLTSLPDRYIKPSPORQRTIIDHOEVADINPIIDDS	

	repeat_region	11962..11996 /fp_family="GA)n"
	misc_feature	12824..12836 /note="exon predicted by xgall, quality marginal.shadowexon"
	misc_feature	/complement(13367..13468) /note="exon predicted by xgall, quality marginal"
	misc_feature	/complement(13742..13788) /note="exon predicted by xgall, quality marginal"
	mRNA	/complement(join(13822..14179,14461..14398,14502..14582,14676..14861,14941..15033,15124..15190,15272..15361,15662..16380)) /gene="F9P8.4"
	gene	/complement(13822..16380) /gene="F9P8.4" /note="identical to omega-3 fatty acid desaturase, chloroplast precursor GB:P46310 [Chloroplast Arabidopsis thaliana]"
CDS		/complement(join(13988..14179,14461..14398,14502..14582,14676..14861,14941..15033,15124..15190,15272..15361,15662..16155)) /gene="F9P8.4" /codon_start=1 /product="Omega-3 fatty acid desaturase, chloroplast precursor" /protein_id="AAFO1508.1" /db_xref="GI:6016681" /translation="MANIVLSEGCIRPLEPIITTPRSNFLSNNNKFRPSLSSSKITGS SSPLVGLNSRDGFTRMNALNVSTPIPTLFEESSPLEEDNKGRFGDGPAPPFNLADIR AALPHCWKNPKMKSLSYVDVAIVFAIAGAAYLNMTIWPVLIAQGTWVALFYR LCHDGHSFENDPRINSYGHLHSHILVPYIGMRISHRTHOHGHGVDESDWHPPR SKIXNTIDKRPTRREFRTLPLYMLAYPEYLMAISPCKSGSHYHPDSDFLPKERKDVL
Query Match	8.0%; Score 84.8; DB 8; Length 101284;	
Best Local Similarity	54.9%; Pred. No. 1,7e+11;	
Matches 167; Conservative	0; Mismatches 137; Indels 0; Gaps 0;	
OY	350 tctctggacactccttgccctgagctgaggaaagctgaaccatttcagtagcgtc 409	
Dd	44936 TCTTTGGTCAGATTCCATCTTGCATTGGGAACCTTGCTCGCTCACTATCTACTC 44877	
OY	410 acaaaacaacatcaagaactataaccttcogaacttggaaatctgaagaatcatca 469	
Dd	44876 CTTTCAACAATTTCGTGGTGAAATCCCATCTCTTTGGCAATCTTAACCACCTGATG 44817	
OY	470 gcttgagatcgtgtacaacaacatcttaagaagatagtcccactcttttggagaattta 529	
Dd	44816 TCTTACAGGTGTGATTCCAATAAGCTCAGTGGAAATGTTCCCATCTCACGTGAAATTGA 44757	
OY	530 agtctcgatctttttaagcgttaatgatgaacacgattgacggggccaatccctatagcac 589	
Dd	44756 CGAAGTGTGCTGCTTATTACTTTCGCCAACAAATCAGTTACAGGACGATCTCAATAATACA 44697	
OY	590 tcactgcaatccccaaagccttaaatgttgttagtgcctaagaatatgatttgtgtggaacca 649	
Dd	44656 TCAATTACTATCAAACTGATGAGACTTTGAAGCAAGTAAACGCTTTTCACTGACGATC 44637	
OY	650 tccc 653	
Dd	44636 TCCC 44633	
RESULT 12		
AC008017	116944 bp DNA PLN 18-SEP-1999	
LOCUS	Arabidopsis thaliana chromosome I BAC F3N23 genomic sequence,	
DEFINITION	complete sequence.	



AC008017  
VERSION AC008017.2 GI:5487890  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsids.  
REFERENCE  
AUTHORS 1 (bases 1 to 116944)  
Fedorispiet,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Chen,S., Harman,P.,  
Hicks,R., Huerta,M., Mason,S., Siepel,J., Zimmerman,M., Buehler,E.,  
Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J.,  
Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
JOURNAL Unpublished  
2 (bases 1 to 116944)  
Fedorispiet,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Chen,S., Harman,P.,  
Hicks,R., Huerta,M., Mason,S., Siepel,J., Zimmerman,M., Buehler,E.,  
Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J.,  
Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
JOURNAL Submitted (10-JUN-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
REFERENCE  
AUTHORS 3 (bases 1 to 116944)  
Fedorispiet,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altifli,H., Araujo,R., Hulzar,L., Rowley,D., Chen,S., Harman,P.,  
Hicks,R., Huerta,M., Mason,S., Siepel,J., Zimmerman,M., Buehler,E.,  
Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J.,  
Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
JOURNAL Submitted (15-JUL-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
TITLE  
JOURNAL Direct Submission  
Submitted (18-SEP-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
COMMENT  
FEATURES  
source  
On Jul 15, 1999 this sequence version replaced gi:5441915.  
location/Qualifiers  
1.116944  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F3N23"  
complement(1333..2966)  
/gene="F3N23.1"  
complement(join(1333..2271,3355..2966))  
/note="81% identical to threonine synthase [Arabidopsis  
thaliana] (gi|4914408). Location of ests GBG589 3'  
(gb|f15267) and GBG5175 (gb|226034)."  
/codon\_start=1  
/product="Putative threonine synthase"  
/protein\_id="A053628.1"  
/db\_xref="GI:5903070"  
/translation="MASFLPHATYFSPHSSETSLKPSASAFYVCTASAPVPPQT  
PKRRSPDENIDEARRRHOLQNSARYPPNAPPSSTESYSLDEIYRSOSGLL  
DNOHDAALKRYDGEFWRNLFDRNGKTNWPGSSGYSKKEWTLPEIDDDIVSAFEG  
NSLFWAERFGKQYLDQNDLWVAHCGISHTGSKDLSVLSQVNRKLKMKKPFYIGV

GCASGDTISALSAVCASGIPSIYVFLPADKISMAQLVOPIANGAFVLSDTDEGCM  
HIREVTAELPIYLIANSLSRLGCOATIAIILQOFMNOVDVIVGCGNATYAF  
YXGFHCKRELGLVDRIPLVYCAQANANPDIYHRSYSGRKEPNLKATFTFASAIQ  
DPSIDIRATYALKSKNGIYEZATEEBLNDATLADSTGMCITCPHGVALTLMLTKRS  
GYVAGDRTVVVSTAHGLKFTQSKIDYHSKNIKENACRLANPAAVAKKRSYMDYLK  
EYLSKNDK"  
6030..7079  
/gene="F3N23.2"  
6030..7079  
/note="Unknown protein: Location of ests 205D1217  
(gb|H7203) and 205D12XP 3' (gb|AA603559)."  
/codon\_start=1  
/protein\_id="A053629.1"  
/db\_xref="GI:5903071"  
/translation="MTGIAEESAEAOETHLPADIMWEMDKSFPVIGALFSGVSA  
LPVAVLAKTRDOVCHSOGSCITATITVNRHBEGLRGYRFGISLMGTIPARALVYTA  
EYKRSVGSAAVSLGLEAKAAVANAAGLSAAADLVWTPVIVVQSRMLVQSGAG  
LVNASCNVNGEDFAFRKIVRADPGKLGFGISILYAPSNVAMASVVAQRMW  
GGIGCYVCKDESGNNTMKPDSKTIAMAVGVSAALAGSVSALITPDLTKTRLO  
VLGDEGSNNKRGKPSIGQTVNRNLVREGGWACYRGIGRCASMSMSATMTITTEFL  
KRLSADKHGFGYSK5"  
complement(8272..9661)  
/gene="F3N23.3"  
complement(join(8272..8658,8746..8916,9005..9079,  
9161..9252,9364..9661))  
/gene="F3N23.3"  
/note="99% identical to Transcription factor [Arabidopsis  
thaliana] (gi|2398525)."  
/codon\_start=1  
/product="Transcription Factor"  
/protein\_id="A053630.1"  
/db\_xref="GI:5903072"  
/translation="MAHOMLNKKSATHTSLPILNLTISMGVPTDVSANRGSASEL  
SLVDSRPHGICOTKQISPODSSSTSGSYEPAVSSGDPNRSROISFAKSGSE  
ITRKRGFASHPKGSSTGTPNITFAPAQNFSHYADHPYGLIATITIPAPTCNPO  
MVNIPGRVPLPALEETDPEVNAKQYHAIMRQRAKLDKQNLKARKPIHES  
RHYHAKLRPGSGRFLNTRKLLQESDQAAARQEDKIGQOVNRTNNSHREANHLQ  
NNKRSSTWTSKSDITSVSDGADIFGHTERFQSGFPPIINRAMLVHGSNDHMGDDMH  
HFSVHT"  
complement(12076..16057)  
/gene="F3N23.4"  
complement(join(12076..12519,12613..13851,13955..14140,  
14368..15451,15672..16057))  
/gene="F3N23.4"  
/note="Similar to disease resistance proteins RPP1-Wsb  
(gi|3860165), RPP1-Wsc (gi|3860167) and RPP1-Wsa  
(gi|3860163) [Arabidopsis thaliana]."  
/codon\_start=1  
/product="Similar to disease resistance proteins"  
/protein\_id="A053631.1"  
/translation="MASSSSSATRLRHVDVLFSPFGVDRQRTIVSHLYVALRNNGV  
LTPKDRKLEIGDITADGIYKAIQTSWPAVYILSENAVSTMCLEELRLMOIHSEIOI  
KVLPITFGVPSDVRIOEGSFATAPSVDEADIAIYVGGISRLRPMKSTDLINVG  
MEAHMKMTLLNLNGCEDEVHMLGIWGGIGSTAKCLIDYRSQFPAHCFLEWVS  
KGYDIKHLQKLESHLIIYRDDEVLMSGASQOPIKRLKHQKVFVLDVAVDEQDLS  
LAKPDWFGEGSSRIIITTRDGLNCSGVAIIEVCLDDKDLQVFKLALGAGRPS  
DGFEOLFIRASRLAHGPPSALVAFASLSIIVADIMDEEDLALLETPEPKNOVIELRA  
SYGDIQDYDKTVFLVAVACFRNGHLIRYIRALNCARINHLAAKLVINISDGCISM  
HILVOTGREIVROESMWRPSKORFLMDPTEIHVYVDSNTHLGGNSNQILSDDYVL  
SRNLKLMWAVYPTLIPLEFRPHITLISLRISKLNSLMDGKLPNLRILVDYDQ  
NLRELPELSTAVNPELIELSCSLVQIPEISIRKLRLKLNMYCGELGCVLIDVLSQ  
EASLSRWGLKRIILNPHSGATLSLTDLAIVQKIFIKLSGSGTDHLSFSQVQTA  
HOSVTHLNGSFGFKSLDKRFSYRDLADPFCISFADPCLTEKLINTNEDLPE  
DIOLOLITLIDIGNDFYVLPISMGQIAMLKTLISNCRRLKALPOLSOVERVLEIS  
CVKLSGLMGLTAGRYVLDPCYKCKSLDSLMGIIISVEKSAQRNLELISLENKS  
LVLSSELSHFTKLTIVDLSLFEFRRLPYSIRLSFMRULYLNCKIRISLIDPESL  
KYLVAHGESLMEHYNESNSHFNHLDIFSCHISLSECDIVRODMNEYSQOAFRLVC  
ITKYSIASITNNMRTSWREPRIKLPIKAPKIVGFVQVIMVVCERCFPHLOPAPFVNC  
WDCSGEGLDRIILKPNLYOSSEEMNNPNPYKWHLVIVQIPTGISLSEIDVQCES  
HIOVPEBGDEIILICGVHNGFVK"  
16543..17907



CDS  
/gene="F3N23.5"  
join(16543..16961,17058..17907)  
/gene="F3N23.5"  
/note="Similar to N-terminal half of downy mildew  
resistance protein RPP5 (Arabidopsis thaliana)  
(gi12109275)."  
/codon\_start=1  
/product="Similar to part of downy mildew resistance  
protein RPP5"  
/protein\_id="AAD55632.1"  
/db\_xref="gi:5903074"  
/translation="MSSSSKFEVPLSPCEDSPKTFVSVLDRLKQDITNPKDSF  
LAPPELAEVGEIPDDKIQBWRVSLKLTINIPALDSRWSNADITELANIELISNR  
KPLTAAGDVLGMDROMOTLXKLDFKAAEVRILGIMGPGIGIKTTIARAYEELIS  
NFKVHVDKAEKICHODRDLKTEKGTQGLDVGIDKIKTFGRKGLIVDGD  
NIKOLKEIYLAHAFIPGSRVIFVQDNNLVESGEVHAEVOSLARDEALQLESHA  
PQOHPPTFESLSRAVHISGFLPLTKIIGSSLRGKDEERWEKEIQLEGDQEKAI  
MEITSRRYRARKKEDEKIKTSFLLSD"

gene  
complement(19050..23350)  
/gene="F3N23.6"  
complement(join(19050..19667,19764..21344,21686..22799,  
22887..23350))  
/gene="F3N23.6"  
/note="Similar to entire sequence of downy mildew  
resistance protein RPP5 (Arabidopsis thaliana)  
(gi12109275)."  
/codon\_start=1  
/product="Similar to downy mildew resistance protein RPP5"  
/protein\_id="AAD55633.1"  
/db\_xref="gi:5903075"  
/translation="MSSSSSPIMRVDFLSRCHDTRKNTIHLKQYDKGVYER  
DKRLKLGDSISEEISRAIONSTVALITISEYASSWCIDLNVMDLHKNTIKYV  
PIFYGDPSHVHNOTGSPTEFDYODSKMNPYRRELTOIALAGDEPCEDEAS  
MIEIYKDISKULLIMQPYDESDIVGNANHERLSPILSMSENVRIKIGWGGIG  
KITIAKLTDSQSGPACFTLVNSKTRKAGVSSWFKLSRLITTRKGLNLY  
KIGPDIKARFGCRVAVVLDVDMQMAHAPQSSWFKGSLITTRKGLNLY  
GVATVYECMDNDALQLENQAFKGLPSELEKELISASWAGLPVALIAYGL  
FVRMTSLKEMDNLALCFLEAPDESVMELKISYDELETPNEDLVHACFNEPVR  
RATLLDGVLOGGLKILAEKSLIETASGYIMHNVOTQARAIYQDSORRRG  
RGVLMNPYEIYELKRNTPODLVEIILHRSNTSPKRETVVKALNRSMLTMYLVN  
ILALFPEVYEGMNPALRLDSDSENLEQDPSMNVNIELITOCCKRLKTPES  
ISYLRLLTLDVSYCEBLASTITTELNRSGQNLILYSGKEVERSTANISIGNH  
IQMFWLDGVNDLSFTTEQQGPKLTKREKQAPBELKREQQEPKRTLLCGFSL  
MRKGRVYKATSEFLDHMMQDQALPNOQLAFSTRTQALQFLPEPHQSEVYKVA  
OGKSOPTSKFHEFTSVDIRFSYSSDGLCFSLSPCYKVELILNLTAKVYD  
CGKLEKLDMSGNDEPTELPMNOLPRLKYSFNRCCRLPALVYDLETIKSGCI  
NOSILELSYARDGGRFOWLEIWDGCKSISITIDOLRHRTIKSYLDLSHREKLP  
SSIEVSSLRITCLNCKRKIKISIBGIPICLKSLYAHGCEILFETVSLPNSVKHLDS  
HCGKLRDHLINQPLNBSGENSEESLGRFAFFGTGVSFTFHIDKSLTIDLPQIMP  
SPKLGFDAVCVYACERPDIQFSPFSVDMWGVRYFCLYLKRPDHFSTDEDEE

CDS  
/gene="F3N23.5"  
join(16543..16961,17058..17907)  
/gene="F3N23.5"  
/note="Similar to N-terminal half of downy mildew  
resistance protein RPP5 (Arabidopsis thaliana)  
(gi12109275)."  
/codon\_start=1  
/product="Similar to part of downy mildew resistance  
protein RPP5"  
/protein\_id="AAD55632.1"  
/db\_xref="gi:5903074"  
/translation="MSSSSKFEVPLSPCEDSPKTFVSVLDRLKQDITNPKDSF  
LAPPELAEVGEIPDDKIQBWRVSLKLTINIPALDSRWSNADITELANIELISNR  
KPLTAAGDVLGMDROMOTLXKLDFKAAEVRILGIMGPGIGIKTTIARAYEELIS  
NFKVHVDKAEKICHODRDLKTEKGTQGLDVGIDKIKTFGRKGLIVDGD  
NIKOLKEIYLAHAFIPGSRVIFVQDNNLVESGEVHAEVOSLARDEALQLESHA  
PQOHPPTFESLSRAVHISGFLPLTKIIGSSLRGKDEERWEKEIQLEGDQEKAI  
MEITSRRYRARKKEDEKIKTSFLLSD"

Query Match 7.7%; Score 81.6; DB 8; Length 116944;  
Best Local Similarity 49.5%; Pred. No. 1.1e-10;  
Matches 210; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 263 caactctgttaactctctgtacatggtcagatcctcgttaaccaagaacgcgtca 322  
DB 84573 CATGTGAAGCCACATGTAAGTGGTGTATGCTGTGATGATTCTTAAGAAAGTTA 84582  
QY 323 ctctgttggaatttggaataacacctctcgtgaacctcttgagctgagctggaagc 382  
DB 84583 CGTCTCAACTTCAACGGGTTACAGGTTTTCAGGCCAATTTGGGCTCGAATAAGGCAAC 84642  
QY 383 ttgaacattacagatctcagagctctacaacaaacaacatccaaggaactacacctcg 442  
DB 84643 TCAAAAGCTTAAGATATGATATGAGTTCTTACAAATTTCTCTGGGATATACCTTCCA 84702  
QY 443 aacttgaaatctgaagaatctcatcagcttgagctctgtacaacaacaactcttaacagga 502  
DB 84703 GTTTAGGGAATGTAAGCTCACTTGTGTACATTTGTTGTGTGAATAAGTTTCTCTGGA 84762  
QY 503 tagttccactctcttggaataaattgaagctctcgtgctctttaagcgcttaatgacaacc 562

DB 84763 AGTCCCATCATCTCCGCTACCTTGAAGACCTTGCCGATTTGTATCTTACAGCAACT 84822  
QY 563 gattcaggaggaacatccctagagacactcagatcgaatcccaagcttaagtgtgatg 622  
DB 84823 CTTTACTGCGGAATTCCTTAATCTTGTTCGATTCGCGGTGTAATATCTGATG 84882  
QY 623 tctcaagcaatgatttggtggaacatcccaacaacaggaaccttgctcacatccct 662  
DB 84883 TTGAGCATTAACAACCTTAACCGGGTTAATCCCTCAAAATGTTGTAAGCTTAAGCACTTC 84942  
QY 683 taca 686  
DB 84943 TTCA 84946

RESULT 13  
INU77888  
LOCUS  
DEFINITION  
INU77888 5033 bp DNA PLN 23-NOV-1996  
cds.  
IPomoea n1l receptor-like protein kinase (lnrpkl) gene, complete  
ACCESSION  
U77888  
VERSION  
U77888.1 GI:1684912  
KEYWORDS  
Japanese morning glory.  
SOURCE  
IPomoea n1l  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Asteridae; euasterids I; Solanales; Convolvulaceae;  
IPomoea.  
REFERENCE  
1 (bases 1 to 5033)  
Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajasevan,M.S.  
Identification and Preliminary Characterization of an Unusual  
Leucine-rich Repeat Receptor-like protein kinase from Morning Glory  
(IPomoea n1l)  
JOURNAL  
Unpublished  
AUTHORS  
2 (bases 1 to 5033)  
Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajasevan,M.S.  
AUTHORS  
Bassett,C.L., Cohen,R.A., Nickerson,M.L. and Rajasevan,M.S.  
TITLE  
Submitted (12-NOV-1996) USDA, ARS, Appalachian Fruit Research  
JOURNAL  
Station, 45 Wiltshire Road, Kearneysville, WV 25430, USA  
FEATURES  
Source  
1..5033  
/organism="IPomoea n1l"  
/cultivar="Violet"  
/db\_xref="taxon:35883"  
/tissue.type="Young Leaf"  
237..251  
681..695  
1144..4566  
/gene="lnrpkl1"  
join(1144..4108,4202..4566)  
/note="leucine-rich repeat; protein kinase homolog"  
/codon\_start=1  
/product="receptor-like protein kinase"  
/protein\_id="AB36558.1"  
/db\_xref="gi:1684913"  
/translation="MKVAVNFTFLSLSTSSIVAAFLNSDGAALLSLRHWTPIPSD  
ITQSNWASDTPSCSWLGVCDRQRFQFPLNLSYGISGEPSIELSHKRYLVLSGN  
GGFGSPQGLGSLHEIDLSNSGNTGIDTGAALONRNLSPNSLIGPPESL  
LSIPHELYFTFGNGLNGSIPSNNGSELTTLTDGALDNOSGAPPSLGTITPOLLY  
LNDNNVGLTPVTNNLNNVYLDVRRNSLYGALPIDPFSCKQIDTISLNNPTGGL  
PRLGNCSTLRFGAFCALSGPLPSCFGQTKLDTYLLAGNHSGRIPEDLRCKSM  
IDIQDQNLBEELPELGMLSQQLYHLVYNNLSGEVPSIMKIOSLSQLOYNL  
SEGLPYDMEIKQIVSLALYNNHFTGVYLPDQGANSSLYEVDLIRNFTGHIIPPLCS  
OKRLRLALGYNYLEGSVPDUGGCSLERLILENNILRGGLPDEVKQMLPEPDLGS  
NNFTGPIPSLGNLKNVAIYLSNOLSGSLPPLGSLVYKLEHLNLSHNTLKGLPSE  
LNCNHLGSLDSHNLNGSLSPITGLSTELTKLSIGENSGSGIPSPISFOSKMLNL  
OIGNVLAGDIPPGALQALASLNSLSKMLGQLPIDGSKLHELEPDVSHNLSGLT  
RYLSTIQSTLFTINISHNLPSCGVPSPLKFLNNSPTFSGNSDLCLNCPADGLACPS  
SLRPNMOSNTNGKGLSTLGIAMVIGALLFTICLPLFSAFLELRCKKSVQBIATISA  
QGGDSLNLNKVLEANTENDRYVIGKAGHTYRATLSPDKVAVKXVLVTGIRNGSY



SVKREITIGKVRHNLKLEEFWLRKREXGLILTYMENGSLHDIHEFTNPKPLDMS  
TRHNIATGTAHGLAYHFDGDAIVHRDIPKPNILDLDEPHIDFIALKLDQSAI  
SIPSTVOCITGYMAEPENAFITVKRESPEYSGVLLLETRKALDPSNGEDIV  
GWRSVWOTGEIOLKIVDPSLDELIDSSVMEQVTEALSLARCAEKEVDKRPFTMDV  
VWOLFWSRSYSSSVRNRSK"

BASE COUNT 1365 a 1060 c 1017 g 1591 t  
ORIGIN

Query Match 7.7%; Score 81.4; DB 8; Length 5033;  
Best Local Similarity 49.9%; Pred. No. 1.1e-10;  
Matches 205; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 249 ccagagctggagatccaaactctgttaactcctgttacctgtccatcgtcgtaca 308  
DB 1281 CCAGAGCTGGAGATCCCAATTCACCTCTGTCTTCAAGCTGGAGATGAAGTGA 1340  
QY 309 agacacacggctacactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 368  
DB 1341 GAGACAAATTTGTTGATCTGTAACCTCTCTCTATGAGATCTCAGGCAATTCG 1400  
QY 369 tgaagctggagagcttgaaactttacagatctagagcttacaacacacacacag 428  
DB 1401 CGAATATCTCGCATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTG 1460  
QY 429 aacatccttcgaacttgaaatctgaagaatctcatcagcttgatctgtacacaa 488  
DB 1461 CTCATATCTCTCCAGATGAGCAATTCAGTCTTCTTGAACATGATGATGCTCC 1520  
QY 489 caatctacagagatgtccactctcttgagaaatgaagctctgtctgtctttttag 548  
DB 1521 CAGCTTACGTGTAATTCCTGACACCTTGAGCTTTCAGATTTAAGAAATTTAG 1580  
QY 549 gcttaagacacacgcatgtgagcggtccatccctagagacactcgtcgaatccca 608  
DB 1581 CCTGTCTTAAATCTCTGATGGCCATTCCTGAGTCTTCAATTCACATTTT 1640  
QY 609 taaagctgtatgtctcaagcaatgtttgtgtgagacacacacacacacacac 659  
DB 1641 AGAAACGTGTTATTTCACTGCAATGCTTAAATGTTCAATCCCTTCAAA 1691

RESULT 14  
AC021198 81513 bp DNA 09-FEB-2000  
LOCUS Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete  
DEFINITION  
AC021198  
VERSION Arabidopsis thaliana chromosome 1 BAC F14D7 sequence, complete  
AC021198.2 GI:6957696  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsis.  
REFERENCE  
AUTHORS 1 (bases 1 to 81513)  
Liu, S.X., Sakano, H., Yu, G., Lee, J.M., Lenz, C., Pham, P., Toriumi, M.,  
Chin, C., Chlou, J., Choi, E., Chung, M., Gonzalez, A., Hwang, B.,  
Liu, A., Vayberg, M., Altai, H., Brooks, S., Buehler, E., Cho, O.,  
Conn, L., Conway, A.B., Hansen, N.F., Johnson-Hopson, C., Khan, S.,  
Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,  
Southwick, A., Davis, R.W., Ecker, J.R., Federspiel, N.A. and  
Theologis, A.  
TITLE Arabidopsis thaliana chromosome 1 BAC F14D7 sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 81513)  
Theologis, A.  
AUTHORS Direct Submission  
TITLE Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan  
JOURNAL Street, Albany, CA 94710, USA  
REFERENCE 3 (bases 1 to 81513)  
Theologis, A.

TITLE Direct Submission  
JOURNAL Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
COMMENT The sequence is of BAC F14D7 from Arabidopsis thaliana chromosome  
1. The sequence does not represent the sequence of the entire  
insert of this clone. It is shorter by 6954 bp because we submit  
only the unique sequence of the clone. However, in order to  
facilitate the joining of overlapping clones in the future for  
creation of larger contigs, we provide small overlaps (200 bp)  
between overlapping submitted clones. The 5' end of this sequence  
overlaps by 200 bp to the 3' end of the sequence of the clone  
F1504.

FEATURES  
Source Location/Qualifiers  
1. 81513  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F14D7"

BASE COUNT 27779 a 13814 c 13868 g 26052 t  
ORIGIN

Query Match 7.5%; Score 79.8; DB 50; Length 81513;  
Best Local Similarity 53.7%; Pred. No. 3.2e-10;  
Matches 165; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 347 accctctgagacatcttgccgctgagcttggaagcttgaacattacagatcagagc 406  
DB 2466 ATCTACCTGGTGGCATTCGCCGGAACATGAGCAACATGATGATGATGATGAGT 2525  
QY 407 tctacaaaacacatccaaagacataccttcgagactggaatctgaagacatcc 466  
DB 2526 TGAGTACACAACTTACCGGTTCAATTCCTTCTTACGAAATCTCAAAATTTGA 2585  
QY 467 tcaagcttgatctgacacacacatctcagagatgtccactctcttggaat 526  
DB 2586 CTATCTTACTTCACTTACGAGAACTATTTAACTGTGTCATTCGCCGGAACATGAGCA 2645  
QY 527 tgaagctctgtctcttttaagcttaagacacacgcatgaggggacatccctag 586  
DB 2646 TGAATTCATGATTTGATTTACGTTGAATTAACAACAACTTACCGGTTCAATTCCTT 2705  
QY 587 caatcagcgaatcccaagccttaagcttgatgtctcaagcaatgtgtgaa 646  
DB 2706 CCTTCGGAATCTCAAGAACTTGAAGTATCTTATCTTACCTGAATTTTAACTGCTG 2765  
QY 647 caatccc 653  
DB 2766 TCATTC 2772

RESULT 15  
ATAC011765 134402 bp DNA 08-DEC-1999  
LOCUS Arabidopsis thaliana chromosome 1 BAC F1M20 genomic sequence,  
DEFINITION complete sequence.  
AC011765  
VERSION Arabidopsis thaliana chromosome 1 BAC F1M20 genomic sequence,  
AC011765.4 GI:6539234  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
Arabidopsis.  
REFERENCE  
AUTHORS 1 (bases 1 to 134402)  
Lin, X., Kaul, S., Town, C.D., Benito, M., Creasy, T.H., Haas, B.,  
Rommig, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E.,  
Bowman, C.L., White, O., Niernman, W.C. and Fraser, C.M.  
TITLE Arabidopsis thaliana chromosome 1 BAC F1M20 genomic sequence  
JOURNAL Unpublished







```
misc_feature      /rpt_family="POLY_A"
9791..9912
/Note="exon predicted by xgrail, quality marginal"
misc_feature      /note="exon predicted by xgrail, quality good"
10574..10662
<10992..>14231
/feature="F1M20.4"
<10992..>14231
/feature="F1M20.4"
gene             /note="similar to brassinosteroid insensitive 1
GB:AC49810 (putative receptor protein kinase); contains
Pfam profiles: PF00560 Leucine Rich Repeat (17 repeats),
PF0069 Eukaryotic protein kinase domain"
10992..14231
CDS              /gene="F1M20.4"
/codon_start=1
/product="putative receptor protein kinase"
/protein_id="AF15908.1"
/db_xref="GI:6539238"
/translation="MAAIVAGDSDSDREVLLSLKSYLESRRPNRGLYTEMKENQ
DYVQMPGICTPQRSRVGINTDSTISGPLEKNSATLETLYDLSRNTGEIPD
DLSRCHNLKHLNLSHNLLEGSLPGLSNLEVLIDLNKRTGDIQSPFLCNSLYA
NLSTNNTGRIDIDFNGCRNLKYVDFSSNRFGEVWTGRLVEFSVADNHLNGNLSA
SMFRGCTQLMDLDSGNAFGEPPGVSNCONLVNLMGNKFTGNIPAEIGSISLK
GLYGNNTFSRDIPELTNLNTLVFLDLSNKRFGDIOEIFGRFTQKYLVLHANSYV
GGINSSNIIKLPLNLSRLDGLNNSGOLPTEISOISLKEFLIAYNFGDIPQEXGN
MPGIALDLSFNRLTSGISIPSPKLTSLIMTMANNSLSGEIPREIGNCTSLMENVY
NNQISGRFPELIRMGSNRPPTPEVNRQNDKRIANGSCILAKRWIRPADPEPPNY
AIIKRSKRSLSMDVHLKVGIFPVCSNGSTVRLKISATIQLSGNKSGEIPASISQ
DLSTLHGFNEFEGKLPPEIGQLPLAFNLTRNNSGEIPQELGNLKLQNLDSLFSN
NFGNPFSTSLNDLNLSEKFNISYNPFSIGAIPTTGQVATFDKDSFLGNPLRPFEN
OSGNNTKISNOVLGNRPRTLLIWLISALALAFIACLVSGIVLVKASREAEIDL
LDGSKTRHDMTSSSGSPWLSGKIKYIRLDSKSTFYADILKATNSSEERYVGRGY
GTYVRGLVPGRERYAVKYLORBESEAKEPRAMEVLSANRQDMAHPNLVRLYGCIL
DSEKTIYHEIMGGSLBELITDTKTLQMKKRIDINIVARGLYFLHHCYPSIYHRD
VKASNVLLDRHGNAKVDFGLARLNLVNGDSHVSIVLAGTIGYVAPETGQWQATTIKGD
```

```
Query Match      7.5%: Score 79.2; DB 8; Length 134402;
Best Local Similarity 52.8%; Pred. No. 4,7e-10;
Matches 171; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
```

```
Qy 329 tggattgggaatcaaacctctcggacatcttgcctgagcttggagaacttgaac 388
Db 12043 TAGATTGGGCTACAACTATTCTCGGACAGTTCACCTACTGAATTTCTCAGATACAGA 12102
Qy 389 attcaagatcctagagctctcaaaaacaacatccagaactatcctccgaacttg 448
Db 12103 GTTGAAGTTCTTGATTCTTGCTTAATAAATTCAAGTGGCATATACACAGAGATATG 12162
Qy 449 gaaatctgaagaatctcatcagcttgatctgtacaacaacatcttaaggatattc 508
Db 12163 GGAACATGCGGGGCTTCAGACACTTGATCTCTCTTAACAAGCTGACCGTTCGATAC 12222
Qy 509 ccactctcttgggaaatgaagtctctgctcttttcaaggcttaataatgaacacgatga 568
Db 12223 CAGCTTCATTGGGAAATGACCTCTCTTGTGGCTAATGCTGCAACAACCTCTCTAT 12282
Qy 569 cggggccaatccctagagactcactgcactccaagccttaagtgtgtatgtctcaa 628
Db 12283 CAGGAGAAATCCCTCGAAGATGGCAACTGACAAAGCCTTTTGCGTTAAACGTGCGCA 12342
Qy 629 gcaatgatttgttggaaacatcc 652
Db 12343 ACAACCAAGCTCTGTGTAGATTCC 12366
```

Search completed: June 24, 2000, 00:26:34  
Job time: 46549 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2000, 22:53:35 ; Search time 446.21 Seconds  
(without alignments)  
596.029 Million cell updates/sec

Title: US-09-180-798-30

Perfect score: 1063  
Sequence: 1 tcgaccacgcgtccgcagc.....tcgcaaaaaaaaaaaaaaaaaa 1063

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : N\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1063	100.0	1063	1	Arabidopsis thalia
2	912.4	85.8	1106	1	Arabidopsis thalia
3	841.8	79.2	981	1	Arabidopsis thalia
4	769.6	72.4	788	1	Arabidopsis thalia
5	361	34.0	894	1	Arabidopsis thalia
6	211.8	19.9	2089	1	Arabidopsis thalia
7	103.4	9.7	1814	1	Arabidopsis thalia
8	75.4	7.1	3176	1	Arabidopsis thalia
9	75	7.1	5940	1	Arabidopsis thalia
10	71.4	6.7	3573	1	Arabidopsis thalia
11	71.4	6.7	6471	1	Arabidopsis thalia
12	70.8	6.7	4081	1	Arabidopsis thalia
13	70.8	6.7	6695	1	Arabidopsis thalia
14	70.6	6.6	3979	1	Arabidopsis thalia
15	70.6	6.6	3979	1	Arabidopsis thalia
16	70.6	6.6	4123	1	Arabidopsis thalia
17	70.2	6.6	9424	1	Arabidopsis thalia
18	67.8	6.4	3541	1	Arabidopsis thalia
19	66	6.2	8416	1	Arabidopsis thalia
20	65.4	6.2	19639	1	Arabidopsis thalia
21	63.6	6.0	3921	1	Arabidopsis thalia
22	63.6	6.0	13340	1	Arabidopsis thalia
23	61	5.7	3293	1	Arabidopsis thalia
24	60.4	5.7	7204	1	Arabidopsis thalia
25	59.2	5.6	3045	1	Arabidopsis thalia
26	58	5.5	2075	1	Arabidopsis thalia
27	56.4	5.3	2192	1	Arabidopsis thalia
28	55.8	5.2	1554	1	Arabidopsis thalia
29	55.8	5.2	4104	1	Arabidopsis thalia
30	55.4	5.2	1058	1	Arabidopsis thalia
31	55.4	5.2	6256	1	Arabidopsis thalia
32	53.2	5.0	3842	1	Arabidopsis thalia
33	53	5.0	5733	1	Arabidopsis thalia
34	49.8	4.7	3050	1	Arabidopsis thalia

35	49.8	4.7	3089	1	Tomato pathogen re
36	48.4	4.6	2880	1	Tomato Cf-9 cDNA.
37	48.4	4.6	2880	1	Tomato pathogen re
38	48.4	4.6	3905	1	Tomato Cf-9 gene.
39	48.4	4.6	3905	1	Tomato pathogen re
40	41.4	3.9	1117	1	Sequence encoding
41	41.4	3.9	2917	1	Raspberry dtul gen
42	40.6	3.8	792	1	Sequence of varian
43	39.4	3.7	1686	1	Sequence encoding
44	37	3.5	692	1	Enterococcus faeca
45	36.4	3.4	2241	1	Staphylococcus aur

## ALIGNMENTS

RESULT	1	
ID	V06590	standard; cDNA to mRNA; 1063 BP.
AC	V06590;	
DT	03-AUG-1998	(first entry)
DE	Arabidopsis thaliana SERK LRR homologous EST clone.	
KW	receptor kinase; apomixis; apomictic; seeds; production; embryos;	
KW	plant breeding; leucine-rich repeat; ss.	
OS	Arabidopsis thaliana.	
FT	Key	Location/Qualifiers
FT	CDS	106..762
FT		/tag= a
FT		/note= "shows high homology to SERK"
PN	W09743427-A1.	
PD	20-NOV-1997.	
PF	13-MAY-1997; E02443.	
PR	14-MAY-1996; GR-010044.	
PA	(NOVS ) NOVARTIS AG.	
PI	De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;	
DR	WPI, 98-086529/08.	
DR	P-PSDB; W47022.	
PT	Production of apomictic seeds - useful in plant breeding	
PS	Claim 28; Pages 86-88; 123pp; English.	
CC	The sequence is that of an EST clone showing high homology to	
CC	SERK LRR (leucine-rich repeat) sequences.	
SQ	Sequence 1063 BP; 313 A; 242 C; 206 G; 302 T;	
Query Match	100.0%; Score 1063; DB 1; Length 1063;	
Best local similarity	100.0%; Pred. No. 3.6e-287;	
Matches 1063; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	tcgaccacgcgtccgcagcgaacccaatttgccttcctcaactctgtccgaagaattact 60
DB	1	TCGACCCACGCGTCGACGAAACCCCTTAATTTGCTTCCATCTTGTCAGAAATTTACT 60
QY	61	caattccctttagattactctctcttcgacctcgcagatgcacatgscgtcgaac 120
DB	61	CAATTCTCTTAGATTACTCTCTCTTCGACCTCGATAGCTACATGGCTCTGGAAC 120
QY	121	tatcgtgtgagctcttcgcagcttcgttaactcctaacttgaacttgaactcgtgc 180
DB	121	TATGGGGGAGGCTCTTCGACGCTTGTAATCTTAACCTTAGCTTGATTCACCTGGTC 180
QY	181	gaacacacccgcgaagaagatgtctttagcgtcttcgcgcgaagttaacagatccgac 240
DB	181	GAAACCAACTCCGAGGAGATGCTCTTTAGCTCTTCGCCGAGTTTAACATATCCGAC 240
QY	241	catgttcacagagcttggatccaactctgttaactccttgaactgttcatgtcacc 300
DB	241	CATGTTCTCCAGAGCTGGATCCAACTCTTGTAATCCTTAGCTGATTCATGTCAC 300
QY	301	tgtacacaaacacacgcgcactcgtgtggatttgggaattcaactccttgaat 360
DB	301	TGTAAACAAACACGCGCTACTCGTGGATTGGGAATTCAACCTCTCTGACAT 360
QY	361	cttgcgctgagcttgggaagcttgaactttagactagactagactctcaaaaacac 420



```
Db 361 |||||CTTGCGCTTAAGCTGGGAAAGCTTGAACATTTTCAGTATCTAGAGCTTCACAAAAAACAC 420
Qy 421 atccaaagaactaactacccgaacttggaacttgaaatctgaagaatctcaactcagctgtagctg 480
Db 421 ATCCAAAGAACTATACCTTCGCAACTTGAATCTGAAATCTCATACAGCTTGATCTG 480
Qy 481 tacaacaacaactcttaagagtagtcccaactctcttggaataattgaagtctcgtgc 540
Db 481 TACAACAACAATCTTACAGGGAATAGTCCACCTTTGGGAAATGGAAGTCTGTCGTC 540
Qy 541 ttcttaagcttaatgaacaaccgattgacggggccaatcccttagagcaactcactgcaatc 600
Db 541 TTTTATAGGCTTATGACAAACCATTTGACGGGGCAATCCCTTAGAGCACTCATGCAATC 600
Qy 601 ccaagcccttaagttgtagtctcgaagaatgattgtgtggaacaatcccaacaac 660
Db 601 CCAAGCCTTAAAGTTGTGATGTCCTCAAGCAATGATTTGTGGAACAATCCCAACAAC 660
Qy 661 ggaactcttgctacatctctcttaagaacttgagaaacccggagttgagggagccg 720
Db 661 GGACCTTTTGCTCATTCTTACAGAACTTGAAGAACCCGAGGTTGAGGGACCG 720
Qy 721 gaactactcgtcttgcaagctacgacaactgacacctgaaanaattgccaanaactg 780
Db 721 GAATTACTCGGCTTGCAGAGTACGACACTAAGTGCACCTGAAANAATTGGCAAAACCTG 780
Qy 781 aaatgaagaattgggggggtagacctgtgaagaacactccaccacttataataatcaca 840
Db 781 AAAATGAGAAATTTGGGGGGTGAACCTTGAAGAACACTTCAACCTTATCAATATATCA 840
Qy 841 tctactatgtaataagatataatataatgtagtccaaanaaaatgaagaatccgaatcagta 900
Db 841 TCTACTATGTAATAGATATATATATGTAATGCCAAAAAATGAAATCGAATCGTA 900
Qy 901 atatcactcgtgctcaacttgagaacttggaacttgatgtatgtataaattcttaaatgcga 960
Db 901 ATATCACTGCTGCTCAATTGGAACCTTGAAGCTGTGTATGTAATTTCTTAATTCGA 960
Qy 961 ctctcgactgtaataatgctggtgtgtggagatctgagaagtaacattgtattgtagtg 1020
Db 961 CTTTCGGCTACTGTAATGCTGTGTGGAGTCTGAGAAATAACATTTGATTTGGTAGT 1020
Qy 1021 gtatcaagttgtctgccttgctgcgaanaaaaaaaaaaaaaa 1063
Db 1021 GTATCAAGTTGTCTGCTTCTGCAAAAAAATTTTTAAAAA 1063

RESULT 2
ID V06586 standard: cDNA to mRNA: 1106 BP.
AC V06586:
DN 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; epomixis; apomictic; seeds; production; embryos;
RM plant breeding; leucine-rich repeat; ss.
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT 142..798
FT /*tag= a
FT /*note= "shows high homology to SERK"
PN W09743427-A1.
PD 20-NOV-1997.
PF 13-MAY-1997; E02443.
PR 14-MAY-1996; GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;
DR WP1: 98-086529/08.
DR P-PSDB: W47018.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 71-73; 123pp: English.
CC The sequence is that of an EST clone showing high homology to
SERK LRR (leucine-rich repeat) sequences.
```

```
SQ Sequence 1106 BP; 331 A; 258 C; 206 G; 311 T;
Query Match 85.8%; Score 912.4; DB 1; Length 1106;
Best Local Similarity 95.0%; Pred. No. 3,6e-245;
Matches 1011; Conservative 0; Mismatches 36; Indels 17; Gaps 6;
Qy 14 ccgaagaacccctaatattgtcttcctcattc-tgtctgaagaanaatct-----actca 62
Db 39 CCAAGCTAACCTTAAATTTTGTGCTTCTCTCTCTTTTATGAAAATTTTCCCTTACTCTCA 98
Qy 63 aattcctatagatctctctctcctcctcgaactcgaatcgaatcgaatcgaatcgaatcga 122
Db 99 AATCTCTTTGATTTCCCTCTTTAAACCTCGAAGACCTACATGCGCTCTCGAACTA 158
Qy 123 tcggtggagagctcttcgagcttcgtaactccttaacttgcttgattacccggtgcga 182
Db 159 TCGGTGGAGAGCTCTTCGAGCTTGTAACCTTACCTTTGATTCACCTGTCGA 218
Qy 183 agcaaacctcgaagagagatgctcttcttaagctcttcgcgcgcgcgttaacagaatccgaca 242
Db 219 AGCAAACTCCGAAGAGATGCTCTCTACGCTCTTCGCCGAGTTGACAGATCCAGACCA 278
Qy 243 tgtctcgaagagctggatcccaactctgttaactcctgttaactcctgttccatgtcacctg 302
Db 279 TGTCTCCAGAGAGCTGGATCCCAACCTTGTAACTTGTACCTGTACCTGTCATGTCACCTG 338
Qy 303 taaccaagaacacgcgcacatctgtgtgatttgaggatccaaactcctcctgaacatc 362
Db 339 TACCAAGAACACCGCGTACCTGCTGTGGAATTTGGAATTTCAAACTCTCTGTGACATCT 398
Qy 363 tgcgcctgagcttggaagcttgaaacattacagatcctagaagctcctacaaanaaacat 422
Db 399 TCGGCTGAGCTTGGAAGCTTGAACATTTACAGATCTACAGCTTACAAAACAACAT 458
Qy 423 ccaaggaactaactcctcgaacttggaactcgaagaatcctacatcgcttgatcgtga 482
Db 459 CCAAGGAATATATACCTTCCGAATCTGGAATCTGAAGATCTCATCACCTTGATCTGTA 518
Qy 483 caacaacaactcttaagagatagttcccaactcctcttggaagaattgaaagtctcgtgctt 542
Db 519 CAACAACAATCTTAAAGGATGATGTTCCCACTTTCTTGGAANAATTGAAGTCTGAGCTT 578
Qy 543 cttaagcttaatgaacacggattgacggggccaatccctcctagaacactcactgaatccc 602
Db 579 TTTAGGCTTATGACAACCGATGACCGGTCCAACT-CTTAGAGACACTACGCAATGCC 637
Qy 603 aagcc-ttaagttgttgatctcgaagcaatgattgtgtggaacaatcccaacaacg 661
Db 638 AAGCCTTAAAGTTGTGACGCTCAACGCAATGTTGTGGGACAATCCCAACAACG 697
Qy 662 gaccttctgacattcctcttaagaacttgagaacacccggagttgagggagccgg 721
Db 698 GACCTTCTCTACATCTCTTACGAACCTTGAACAACCCGAGATTGAGAGGACCGG 757
Qy 722 aattactcgtctgcaagctcgaactcgaactcgaactcgaactcgaactcgaactcga 781
Db 758 AATTACTCGCTTGCAACTGACACTGACACTGACACTGAAACCTGGAACCACTCGA 817
Qy 782 aaatgaagaattgggggggagcttgtaagaacacttaccacttatcaatcacat 841
Db 818 AATGGAAGAAATTTGGGGGGGACCTTGAAGAACACTTCAACCTTATCAATATCACAT 877
Qy 842 ctactatgaataaagatatatactagtcgaa--aaaaaaatgaagaatcgaatcagt 899
Db 878 CTAATATGTAATAGATATATATATGTAATGTAATGTAATGTAATGTAATGTAATGTA 937
Qy 900 aataatactgblctcaatcgaacttgagagcttgatgtatgtaaaattcctaataatgcg 959
Db 938 AATATCATCTGCTCAATTTGGAACCTTGAAGTCTGAGTCTG-CTGATGTAATTTCTAATTCG 995
Qy 960 acttcgctactgtaataatgctggtgtgagatctctggaagaatcattgtattgtagt 1019
```



Query Match	79.2%;	Score 841.8;	DB 1;	Length 981;
Best Local Similarity	96.1%;	Pred. No. 1.7e-225;		
Matches 885;	Conservative 0;	Mismatches 32;	Indels 4;	Gaps 2;
QY 1020 gytacaaagtgtctgcgtgtctgtctgcgaataaaataaaataaa 1063				
Db 1056 GGTATCAAGTTGTTCTGCTGTGCAAAAAAAAAAAAAAAAAA 1099				
RESULT 3				
V06587				
ID V06587	standard;	CDNA to mRNA;	981 BP.	
AC V06587;				
DT 03-AUG-1998 (first entry)				
KW Arabidopsis thaliana SERK LRR homologous EST clone.				
RK receptor kinases; apomixis; apomictic; seeds; production; embryos;				
KW plant breeding; leucine-rich repeat; ss.				
OS Arabidopsis thaliana.				
FS Key	Location/Qualifiers			
FT FT	/tag= a			
FT FT	104..760			
FT CDS				
PT M09743427-A1.	/note= "shows high homology to SERK"			
PD 20-NOV-1997.				
PR 13-MAY-1997; E02443.				
PR 14-MAY-1996; GS-010044.				
PA (NOVS ) NOVARTIS AG.				
PI De Vries SC, Hecht VEG, Schmidt EDL, Van Holst GJ;				
DR WPJ: 98-086529/08.				
DR P-PSDB: M47019.				
PT Production of apomictic seeds - useful in plant breeding				
PS Claim 26; Pages 75-77; 123pp; English.				
CC The sequence is that of an EST clone showing high homology to				
CC SERK LRR (leucine-rich repeat) sequences.				
SC Sequence 981 BP; 286 A; 236 C; 180 G; 279 T;				
QY 59 ctcaaatctctatagattactctctctctgcagccctgcagatgcacatgctctcgaa 118				
Db 57 CTGAATCTCTTTCGATTCCCTCCTCTTAACCTCGAAGCTCAATGCGCTTCGAA 116				
QY 119 actatcggtgagagctctctgcagcttcgtaactctaaccttgccttgattaccctg 178				
Db 117 ACTATCGGTGGAGCTCTTCGCGAGCTTCCTTAACCTTAACCTTGAATTCACCTGG 176				
QY 179 tcgaagaagaactcgaagaagagatgcctcttaagccttcgcgcgaagttaacatccg 238				
Db 177 TCGAAGCAAACTCGAAGAGATGCTCTTACGCTCTTCGCGGAGTTGACGATCCAG 236				
QY 239 accatgtctctcagagctggagatccaaactctgttaactccttgcactggltccatgtca 298				
Db 237 ACCATGTCCTCCAGAGCTGGATCCAACTTGTAACTCTGTAACTGGTTCAGTCA 296				
QY 299 ccgttaaccagaacacccgctgactcgtgtgtgatttggggaattcaaacctctcggac 358				
Db 297 CCTTAACCAAGAACAAACCGCTACTCGTGTGATTTGGAAATTCAAACCTCTCGGAC 356				
QY 359 atcttgagctcgtgagcttgggaagcttgaacatttaacatctatctagagctctacaataca 418				
Db 357 ATCTTGCGCGCTGAGCTGGGAACTTGAAACATTTCAGATCTAGAGCTGTACAAAAACA 416				
QY 419 acatccaagaactatactctccgaacttggaaacttgaacatctgaatcatcaactgattgac 478				
Db 417 ACATCCCAAGGAACATATACCTTCGAACTTGGAAATCTAAGAAATCTATCAGCTTGATC 476				
QY 479 tgtacaacaacaatcttaacagagatagttcccaactcttttgggaataattgaagctctg 538				
Db 477 TGTACAAACAAATCTTACAGGAGATAGTTCACACTCTTTTGGGAAATTAAGTCTCTGG 536				
QY 539 tcttttaaggcttaatgacaaccgatgaacgggggccaatcccttagagactcaatgaca 598				

Query Match	72.4%	Score 769.6	DB 1	Length 788
Best Local Similarity	98.9%	Pred. No. 2.1e-205		
Matches 775	Conservative 0	Mismatches 9	Indels 0	Gaps 0
Db	5	CCCGAGCGCTCGGGAACATATCGTGGGAGCTTCGAGCTTGTTAAATCCATAACCTTA	64	
Qy	103	caaatgacgtctcgaaactatcggtgtagaagctcttcgacagctcggttaactcaactta	162	
Db	5	CCCGAGCGCTCGGGAACATATCGTGGGAGCTTCGAGCTTGTTAAATCCATAACCTTA	64	
Qy	163	gcttggttaactcggtcgaaacaacccggaagagatgctcttaacgcttcgacgg	222	
Db	65	gcttggttaactcggtcgaaacaacccggaagagatgctctcttttgcgctttgcg	124	
Qy	223	agtttaacaatccggaacatglttccagaagctggatccaactctgttaactctgt	282	



```
Db 125 AGTTTAAACGATCCGACCATGTCTCCAGAGCTGGATCCAACTCTTTTAATCCTTGT 184
Qy 283 accgtgtccatgctcaactcgttaacaaagacacccgctcaactcgtgtggaattgggaat 342
Db 185 ACCGTGTTCCATGTCACCTGTAAACCAACACCGGCTACCTCGTGTGATTTGGGGAAAT 244
Qy 343 tcaaacctctcggaaattcttcgcctgagcttggaaatctgaaacatttaagatctta 402
Db 245 TCAAAACCTCTGGACATCTTGCGCTGAGCTTGGAACTTGAACATTTTACAGATCTTA 304
Qy 403 gaggcttaaaaaaacacataccaagaatactacccctcgaactctgaaatctgaagaat 462
Db 305 GAGCTTACAAAACAAACATCCAGAACTATACCTTCGAACTTGGAAATCTGAAGAAAT 364
Qy 463 ctcatcagcttgatctgtacaacaacaaatcttaacaggaatagttcccaactctttgga 522
Db 365 CTCATAGCTTGATCTGTACAAACAAATCTTACAGGATAGTCCCACTCTTTGGCA 424
Qy 523 aaattgaagctctggtctcttttaagagcttaatgaaacagattgaacgggccaatccct 582
Db 425 AAATTGAAGTCTGTGCTTTTACGGCTTAATGACACCGATTTGACGGGCAATCCCT 484
Qy 583 agagcactcactgcaatcccaagccttaagatgtgtatgtctcaagcaatgatttgt 642
Db 485 AGAGCACTCAGCTCAATCCCAAGCCTTAAAGTGTGATGTCTCAAGCAATGATTTGTGT 544
Qy 643 ggaacaatcccaacaaacggaccttttgcacatctcttaacagaactttgaagaacac 702
Db 545 GGACACATCCCAACAAACGAGACCTTTTGTCAATCTTTCAGAACTTTGAGAAACAC 604
Qy 703 cgaaggttggagggagccgaataactcagctctctgcaagcttaacgaactaacgcaactga 762
Db 605 CCAAGGTTGGAGGAGCGGAGATTAATCTCGCTTCGCAAGCTACGACACTAATGCACTGA 664
Qy 763 aaaaattggcaaaacctgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 822
Db 665 AAAAATTGGCAAAACCTGAAATGAAGAAATGGGGGCTGACCTTTGAAACAACTTCAC 724
Qy 823 actttcaaatatcatcatctactatgataatgataatgataatgataatgataatgataatgataat 882
Db 725 ACTTTATCAATATCAATCTACTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 784
Qy 883 tga 886
Db 785 AAAA 788

RESULT 5
V06589
ID V06589 standard; cDNA to mRNA; 894 BP.
AC V06589;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK LRR homologous EST clone.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT 1..678
FT /*tag= a
FT /note= shows high homology to SERK"
PN W09743427-A1.
PD 20-NOV-1997.
PE 13-MAY-1997: E02443.
PR 14-MAY-1996: GB-010044.
PA (NOVS ) NOVARTIS AG.
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR p-PSDB; W47021.
PT Production of apomictic seeds - useful in plant breeding
PS Claim 28; Pages 83-84; 123pp; English.
CC The sequence is that of an EST clone showing high homology to
CC SERK LRR (leucine-rich repeat) sequences.
SQ Sequence 894 BP; 270 A; 163 C; 176 G; 285 T;
```

```
Query Match 34.0%; Score 361; DB 1; Length 894;
Best Local Similarity 72.8%; Pred. No. 2e-91;
Matches 538; Conservative 0; Mismatches 110; Indels 91; Gaps 2;

Qy 189 ctccgaagagatgagctctcttaacgctctccggaagcttaacagatccggaactgtct 248
Db 15 CTCGAAAGGGAGCGCTCTTACCGCGCTTCGCCGACCTATACATCCAGACAAATGTGT 74
Qy 249 ccagagctggatcccaactctgttaatcctgttacctgggtccatgtaacctgaacca 308
Db 75 TCAAGTTGGATCCAACTCTGTAACTCTGTAACTCTGTAACTCTGTAACTCTGTAACTCT 134
Qy 309 agaaacccgctcaactcgtgtgatttgggaattaaacctctcggacacttgcgc 368
Db 135 ACACCAATCAAGTCACTGCTGTGATTTGGGAAATCAAACTATCTGACATCTAGTACC 194
Qy 369 tgaagcttgggaagcttgaacattacagtaact----- 402
Db 195 TGAACCTGGGAAGCTTGAACATTTACAAATATCTGTATGAAATCATCATCTTTTGCCTT 254
Qy 402 ----- 402
Db 255 TGATTATCTGAAAACATTTACATTATCATCATATATATATATATATATATATATATATATATATAT 314
Qy 402 --- agagctctacaaacaaacacatccaaagaaactataactccgaacttggaaatctga 458
Db 315 TAGTGAACCTCTCAAAAACGAGATTCAGGAACTATACCTTGTGACCTTGGAAATGTGA 374
Qy 459 gaatcactcagcttggatcgtgtacaacaacaaactcttaacaggaatagttcccaactt 518
Db 375 GAGCTATATCAGTTTGGATCTGTACAAACAAATCTCAACGGGAAATATCCATCTTCTT 434
Qy 519 gggaaatgaaatcgtctgtcttcttaacgcttaatgaaacacccgattgacgggccaat 578
Db 435 GGGAAATGGAAGTCACTTGTCTTTTGGCGCTTAAGAAACCGATGACCGGCTCTAT 494
Qy 579 cccctagagcactcactcgaatcccaagccttaagatgtgtatgtctcaagcaatgat 638
Db 495 TCCTAGAAACCTCAACAGTATTTTAACCTTAAAGTTGATGATCTCAAGGAATGATTT 554
Qy 639 gtgtgaaacaatcccaacaaacggaccttctgcaacatcccttaacagaacttgaaga 698
Db 555 GTGTGAACAAATTCAGTAGAAGAGACCTTTGAACAACTTCATGCAAAACTTTGAA 614
Qy 699 caaccgaggttggagggagccgaataactcagctctcgaagctaaagcaactaactgcac 758
Db 615 CAACCTGAGATTGGAGGAGCAGAACTACTAGCTTTCGAGCTATGACACCAATTCAC 674
Qy 759 ctgaaaaaattggcaaaacctgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 818
Db 675 TTTAAAGAAATGTGAAGAACTATTAAGAAAGAA-TGTTAGTGACCTTTGAAGAACTCTG 753
Qy 819 caacccttatcaaatatc 837
Db 734 TACCAAGTGTGTGAATC 752

RESULT 6
V06591
ID V06591 standard; cDNA to mRNA; 2089 BP.
AC V06591;
DE 03-AUG-1998 (first entry)
DE Arabidopsis thaliana SERK gene.
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;
OS Arabidopsis thaliana.
FH Key location/Qualifiers
FT 195..2072
FT /*tag= a
FT /product= SERK protein
PN W09743427-A1.
```



PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GS-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
DR P-PSDB: M47023.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 27; Pages 91-95; 123pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 2089 BP; 568 A; 419 C; 503 G; 599 T;

Query Match 19.9%; Score 211.8; DB 1; Length 2089;  
Best Local Similarity 63.0%; Pred. No. 1.1e-49;  
Matches 327; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 184 gcaaacctcgaagagatgctcttaccgtctccgcgaggttaacagatccgagcaat 243  
DB 270 GCTATTGTGAAGGTGATGCTTGCATACCTTGAAGGTTACTGATTCATCCAAACAT 329  
QY 244 gtctccagagctggagatcaactctgttaacctgtgaacctgtgctcatgtcacctgt 303  
DB 330 GTCTTGACAGAGCTGGAGTCTTACGCTAGTGAATCTTGACATGTTCCATGTCACTTGC 389  
QY 304 aaccaagaacaaccgcgtcaactcgtgtgatttgggaatcaaacctctctgacatctt 363  
DB 390 AACAAACGAAAGAGTGTCTATAGAGTTGATTGGGAAATCGAAGTTATCTGGCATTTTA 449  
QY 364 gcgcctgagcttgggaagcttgaaacctttacagtatcttagagcttcacaaaacacatc 423  
DB 450 GTTCACAGAGCTGGTGTGCTCAAGAAATTTGCAATTTTGAAGCTTTCACATCAACATTA 509  
QY 424 caaggaactataccttcgaacttgaagaatcgaagaatctcatcagcttgaactgtac 483  
DB 510 ACTGGCCCGAATCTCTACTAATCTTGGAATCTGACAAACTTACTGATTTGATCTTTAC 569  
QY 484 aacaacaatttaccagagatagttcccaacttcttgggaanaattgaagttctgtgcttt 543  
DB 570 TTAACACAGCTTCTCCGGTCTCTATTCGGAATCAATTTGGGAAGCTTTCAAAGCTGAATTT 629  
QY 544 ttacggcttaatgacaacgatttagcggggccaatcccttagcagcatcactgtaaccca 603  
DB 630 CTCGGGCTTAACAAACAGTCTACCTGCGGTCATTTCTTATGTCATCTGACCAATATTAAT 689  
QY 604 agacctaaagtgtgtatgtctcaagaacatgttgttgggaacaatcccaacaacgga 663  
DB 690 ACCCTTCAAGTGTAGATCTATCAATAACAGACTCTCTGTTCAAGTCTCTGACATGCG 749  
QY 664 cctttgtctcacatcttaccagaacttggagaacac 702  
DB 750 TCCTTCTACTCTTCAACACCATCATGTTTGTCTATTAAC 788

RESULT 7

V06571  
ID V06571 standard; cDNA; 1814 BP.  
AC V06571;  
DT 03-AUG-1998 (first entry)  
DE Daucus carota SERK gene.  
KW receptor kinase; apomixis; apomictic; seeds; production; embryos;  
KW plant breeding; ss.  
OS Daucus carota.  
FH Key Location/Qualifiers  
FT CDS  
FT 94..1755  
FT /tag= a  
FT /product= SERK protein  
FN MO9743427-A1.  
PD 20-NOV-1997.  
PF 13-MAY-1997; E02443.  
PR 14-MAY-1996; GS-010044.  
PA (NOVS ) NOVARTIS AG.  
PI De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;  
DR WPI: 98-086529/08.  
DR P-PSDB: M47013.  
PT Production of apomictic seeds - useful in plant breeding  
PS Claim 21; Pages 47-51; 123pp; English.  
CC The sequence is that encoding SERK, a putative receptor kinase.  
CC It may be used as part of a method of producing apomictic seeds  
CC comprising: (a) transforming plant material with a nucleotide  
CC sequence encoding a protein which in active form in a cell or  
CC cell membrane renders the cell embryogenic; (b) regenerating  
CC the transformed material into plants or carpel-containing  
CC plant parts; and (c) expressing the sequence in the vicinity  
CC of the embryo sac. The apomictic seeds and embryos thus produced  
CC can be developed into plant progeny. This is useful in plant  
CC breeding programs. Controllable and reproducible apomixis provides  
CC many advantages in plant improvement and cultivar development in  
CC the case that sexual plants are available as crosses with the  
CC apomictic plant. Apomixis provides for true-breeding, seed  
CC propagated hybrids and could shorten and simplify the breeding  
CC process so that selfing and progeny testing to produce and/or  
CC stabilise a desirable gene combination could be eliminated.  
CC Apomixis allows plant breeders to develop cultivars with  
CC specific stable traits for such characteristics as height,  
CC seed and forage quality and maturity.  
SQ Sequence 1814 BP; 530 A; 354 C; 415 G; 515 T;

Query Match 9.7%; Score 103.4; DB 1; Length 1814;  
Best Local Similarity 55.4%; Pred. No. 1.8e-19;  
Matches 200; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 363 tgcgcctgagcttgggaagcttgaacatttacagttactagagctctacaacaaacat 422  
DB 138 TGATGCTTACCTTGACAAATATGGGGTCTTATGACATTGGAGCTTTACAGCAATTAACAT 197  
QY 423 ccaaggaactataccttcgaacttgaagaatcgaagaatccatcagttgagcttga 482  
DB 198 AAGTGGACCAATTTCTTGATATCTTGGGAATCTGCAAAATTTGGTGAAGTGGACCTATA 257  
QY 483 caaacaacatttaccagagatagttcccaacttcttgggaanaattgaagttctgtgctt 542  
DB 258 CATGAATAGCTTCTCTGACCTATACCGGACATTTAGGAAGCTTTCACAAAGCTTAAGATT 317  
QY 543 tttagcgcttaatgacaacgatttagcggggccaatcccttagagacactactgcaatccc 602  
DB 318 CTTCGGCTCAACAAACAGCCCTCTCTGGTCCAAATTTCCAAATGTCATCTGACATATTAATAC 377  
QY 603 aagacctaaagtgtgtatgtctcaagaacatgttgttgggaacaatcccaacaacgga 662  
DB 378 AACCTTCAAGTCTCTGATTTATCAAAACATGCGGCTATCAGACGATCCGATATATG 437  
QY 663 acccttgcacatcttccattacagaacttggagaacaaaccgaggttggaggagacgga 722  
DB 438 CTCATTTCTTGTGTATACACCTATACAGTTTGGCAATAATTTGAATTTATGTGACCCGT 497  
QY 723 a 723



Db 498 A 498

## RESULT 8

T62124 standard: cDNA to mRNA; 3176 BP.

T62124;

10-JUN-1997 (first entry)

Arabidopsis thaliana plant morphogenesis regulatory protein cDNA.

Plant; morphogenesis; regulation; short; stem; alteration;

Inflorescence; extraneous; gene; expression; transformation;

Increase; control; form; length; ds.

Arabidopsis thaliana.

Key Location/Qualifiers

FT cds 51..2981

/note="plant morphogenesis regulatory protein"

J09056387-A.

04-MAR-1997.

24-AUG-1995; 216187.

24-AUG-1995; JP-216187.

(MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.

(CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.

WPI: 97-206629/19.

P-PSDB: M13408.

DNA encoding plant morphogenesis regulatory protein - useful to

yield plants with short stems or altered inflorescence

Claim 1; Pages 6-10; 17pp; Japanese.

The present sequence encodes an Arabidopsis thaliana plant

morphogenesis regulatory protein (MRP), which can be used to yield

a plant with, e.g. short stems or altered inflorescence. The MRP

acts on a plant at a specific site for a specific period, and can

therefore be used to regulate extraneous gene expression in a

plant. The MRP's cDNA or genomic DNA can be used to transform a

CC plant to increase its MRP expression, and therefore control the

CC form (particularly stem length) of the plant.

Sequence 3176 BP; 927 A; 654 C; 659 G; 936 T;

Query Match 7.1%; Score 75.4; DB 1; Length 3176;

Best Local Similarity 52.0%; Pred. No. 1.5e-11;

Matches 169; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 329 tggatttgggaattcaaacctctctggaactctgagcctggagctgggaagctgaac 388

Db 985 TGTATTGACACAGTAACAGCTGACTGTTCAATTCCACTGACCTTGAACATGTCAA 1044

QY 389 attacagatctagagctctacaaacaacatccaaggaactatccttcgaacttg 448

Db 1045 AACGCCATTACTGGAACCTCAATATATATCTACAGGGTCATATACCAACAGACTTG 1104

QY 449 gaaatctgaagaatcctcagcttgatctgtacaaacaacatcctcgaaggaagttc 508

Db 1105 GGAAGCTTACTGACTTGTGTTGATGTGATGAGCCAAACAAATCTGGAAGGACCTATAC 1164

QY 509 ccaactctctgggaataatgaagctcgtgctcttttaacggctgaagcaccgattga 568

Db 1165 CTGATCATCTAGCTCTTGCAACAAATCTAAACAGCTTAAATGATGGAACAAAGTTTA 1224

QY 569 cggggccaatcccttagagcactcactgcaatccaagccttaaaagtgttgatctcaa 628

Db 1225 GTGGACATATACCCGAGCACTTTCAAAAGTAGAAGATGACTTACTTATCTGTCCA 1284

QY 629 gcaatgatttgttggaacatccc 653

Db 1285 GCAACAATATCAAAAGTCCAAATCCC 1309

## RESULT 9

X23526 standard: DNA; 5940 BP.

X23526;

DT 17-JUN-1999 (first entry)

DE O. Longistaminata Xa21 gene family member A2 DNA.

KW Xa21; receptor kinase-like protein; multigene family; RKK; rice;

KW plant disease resistance; cassava; maize; tomato; Xanthomonas; ss.

OS Oryza longistaminata.

PN M09090151-A2.

PD 25-FEB-1999

PF 17-JUL-1998; U14841.

PR 13-AUG-1997; US-910386.

PA (RSGC) UNIV CALIFORNIA.

PI Hubert SH, Richter T, Ronald PC, Song W, Szabo V,

PI Wang G;

DR WPI: 99-204431/17.

PT New RKK polynucleotides and nucleic acid constructs - used for

generating transgenic plants resistant to Xanthomonas

PS Claim 1; Page 52-53; 67pp; English.

CC This invention describes a method for conferring disease resistance in

CC plants. The invention describes the use of novel genes and proteins

CC belonging to the Oryza longistaminata and Oryza sativa receptor

CC kinase-like protein (RKK) Xa21 multigene family. Such genes from

CC cassava, maize and tomato are also described. The genes and proteins can

CC be used for enhancing resistance to Xanthomonas in a plant, preferably in

CC rice or tomato.

SQ Sequence 5940 BP; 1570 A; 1200 C; 1188 G; 1982 T;

Query Match 7.1%; Score 75; DB 1; Length 5940;

Best Local Similarity 52.4%; Pred. No. 2.5e-11;

Matches 165; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 343 tcaaacctctctgacacatctgagccttgagcttggaacttggaacttaccagatccta 402

Db 2406 TCCGACCTGTCGGGATATCTCGCGTCGGGCAACCTGCTTCCCTAGAGACGCTG 2465

QY 403 gagctctcaaaaacaacatccaaggaactaccctccgaacttggaacttgaagaat 462

Db 2466 GACCTCAGGAGAACACACCTGTCGCGCAAGATACCTAGGACAGCAGCTCAGAGG 2525

QY 463 ctacacagcttgagatctctacaaacaacatcttaccaggaagatcttccactcttggga 522

Db 2526 CTCCAAACAACTGTAATCTGAATTTCAACAGCCTATACGGGTGAGATTTCCAGCTTTGGCC 2585

QY 523 aaattgaagctctgctgtcttctttaaagcttaagcaccgattgagggccaatccct 582

Db 2586 AATCTAACAGATCTGCTCGGCTTCTGAGCTGACTTAACATACATGCTGAGCAATCCCT 2645

QY 583 agagacatctacgcgaatcccaagccttaagttgtgtgtcctcaagaatgttgt 642

Db 2646 TCACTCTGCGCAACCTCAACAGCCTCACTGATCTGACTGCTGAAAATATGCTGTCT 2705

QY 643 ggaacaatccaaca 657

Db 2706 AGTTCCATCCCTTCA 2720

## RESULT 10

T06307 standard: cDNA; 3573 BP.

T06307;

DT 14-APR-1996 (first entry)

DE Partial tomato pathogen resistance gene Cf-2.2 cDNA clone.

KW Pathogen resistant; Cf-2.2; tomato; C.fulvum; Avr 4; Avr 9; fungal;

KW leaf mould; variegation; ss.

OS Lycopersicon esculentum.

PN M09531564-A2.

PD 23-NOV-1995.

PF 11-MAY-1995; G01075.

PR 11-MAY-1994; GB-009394.

PR 23-DEC-1994; WO-G02812.

PR 31-MAR-1995; GB-006658.

PR 07-APR-1995; GB-007232.

PA (GATS-) GATSBY CHARITABLE FOUND.

PI Hammond-Kosack KE, Jones DA, Jones JDG;







```

FT      /number= 4
FT      4430. .4528
FT      /tag= f
FT      /number= 5
FT      4642. .4757
FT      /tag= g
FT      /number= 6
FT      4890. .4967
FT      /tag= h
FT      /number= 7
FT      5295. .5803
FT      /tag= i
FT      /number= 8
FT      6197. .6339
FT      /tag= j
FT      /number= 9

MO9743427-A1.
PD      20-NOV-1997.
PE      13-MAY-1997; E02443.
PR      14-MAY-1996; GB-010044.
PA      (NOVS ) NOVARTIS AG.
PI      De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR      WPI; 98-086529/08.
DR      P-PSDB; W47017.
PT      Production of apomictic seeds - useful in plant breeding
PS      Claim 26; Pages 64-67; 123pp; English.
CC      The sequence is that encoding SERK, a putative receptor kinase.
CC      It may be used as part of a method of producing apomictic seeds
CC      comprising: (a) transforming plant material with a nucleotide
CC      sequence encoding a protein which in active form in a cell or
CC      cell membrane renders the cell embryogenic; (b) regenerating
CC      the transformed material into plants or carpel-containing
CC      plant parts; and (c) expressing the sequence in the vicinity
CC      of the embryo sac. The apomictic seeds and embryos thus produced
CC      can be developed into plant progeny. This is useful in plant
CC      breeding programs. Controllable and reproducible apomixis provides
CC      many advantages in plant improvement and cultivar development in
CC      the case that sexual plants are available as crosses with the
CC      apomictic plant. Apomixis provides for true-breeding, seed
CC      propagated hybrids and could shorten and simplify the breeding
CC      process so that selfing and progeny testing to produce and/or
CC      stabilise a desirable gene combination could be eliminated.
CC      Apomixis allows plant breeders to develop cultivars with
CC      specific stable traits for such characteristics as height,
CC      seed and forage quality and maturity.
CC      Sequence 4081 BP; 1120 A; 770 C; 765 G; 1406 T;
SQ

Query Match      6.7%; Score 70.8; DB 1; Length 4081;
Best Local Similarity 69.6%; Pred. No. 3.2e-10;
Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      195 agagagatgctcttcagctcttcgcgcgagtttaacagatccgagcatgttccagag 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1794 AGGTGATGCTTTCATCACTTTAGGGATTACTTACTTCACCAACAACTTTCACAG 1853
QY      255 ctggagatcaacatcttgtaacatcctgtacccgtgtccatgcatcgttaacagaa 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1854 CTGGGATCTTACGCTTAGTGAATCTTGCACATGTTCCATGCTTGCACACAGAA 1913
QY      315 ccgcgtactcgtgtgga 332
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1914 CAGTGTCAATAGAGGTGA 1931

RESULT 13
ID      V06570
AC      V06570;
DT      03-ADG-1998 (first entry)
DE      Daucus carota SERK gene.
KW      receptor kinase; apomixis; apomictic; seeds; production; embryos;
      plant breeding; ds.

```

```

OS      Daucus carota.
PH      Key
FT      CDS
FT      /tag= a
FT      /note= "contains introns"
FT      3731. .3802
FT      /tag= b
FT      /number= 1
FT      3851. .3979
FT      /tag= c
FT      /number= 2
FT      4124. .4211
FT      /tag= d
FT      /number= 3
FT      4284. .4357
FT      /tag= e
FT      /number= 4
FT      4430. .4528
FT      /tag= f
FT      /number= 5
FT      4642. .4757
FT      /tag= g
FT      /number= 6
FT      4890. .4967
FT      /tag= h
FT      /number= 7
FT      5295. .5803
FT      /tag= i
FT      /number= 8
FT      6197. .6339
FT      /tag= j
FT      /number= 9

MO9743427-A1.
PD      20-NOV-1997.
PE      13-MAY-1997; E02443.
PR      14-MAY-1996; GB-010044.
PA      (NOVS ) NOVARTIS AG.
PI      De Vries SC, Hecht VFG, Schmidt EDL, Van Holst GJ;
DR      WPI; 98-086529/08.
DR      P-PSDB; W47013.
PT      Production of apomictic seeds - useful in plant breeding
PS      Claim 21; Pages 40-46; 123pp; English.
CC      The sequence is that encoding SERK, a putative receptor kinase.
CC      It may be used as part of a method of producing apomictic seeds
CC      comprising: (a) transforming plant material with a nucleotide
CC      sequence encoding a protein which in active form in a cell or
CC      cell membrane renders the cell embryogenic; (b) regenerating
CC      the transformed material into plants or carpel-containing
CC      plant parts; and (c) expressing the sequence in the vicinity
CC      of the embryo sac. The apomictic seeds and embryos thus produced
CC      can be developed into plant progeny. This is useful in plant
CC      breeding programs. Controllable and reproducible apomixis provides
CC      many advantages in plant improvement and cultivar development in
CC      the case that sexual plants are available as crosses with the
CC      apomictic plant. Apomixis provides for true-breeding, seed
CC      propagated hybrids and could shorten and simplify the breeding
CC      process so that selfing and progeny testing to produce and/or
CC      stabilise a desirable gene combination could be eliminated.
CC      Apomixis allows plant breeders to develop cultivars with
CC      specific stable traits for such characteristics as height,
CC      seed and forage quality and maturity.
CC      Sequence 6695 BP; 1844 A; 1182 C; 1243 G; 2422 T;
SQ

Query Match      6.7%; Score 70.8; DB 1; Length 6695;
Best Local Similarity 69.6%; Pred. No. 3.9e-10;
Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      195 agagagatgctcttcagctcttcgcgcgagtttaacagatccgagcatgttccagag 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      956 AGGCGATGCTTACACACACTTACGACTAGCTTGCAGAGTCCCAACAACTCTTCAGAG 1015
QY      255 ctggagatcaacatcttgtaacatcctgtacccgtgtccatgcatcgttaacagaa 314

```



Db 1016 CTGGGATCCACACCCCTTGTGACCCCTTGACACATGGTTTCATGTGACATGTAAACAATGAAA 1075  
 |||||  
 Oy 315 ccgcgtcactcgtgtgga 332  
 |||||  
 Db 1076 CAGTGTTATAGAGATGTA 1093  
 |||||

## RESULT 14

1V14518	
ID	V14518 standard; DNA; 3979 BP.
AC	V14518:
DE	20-MAY-1998 (first entry)
DE	CF-5 pathogen resistance gene variant #1.
KM	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW	tomato leaf mould; Phytophthora resistance; ss.
OS	Lycopersicon pimpinellifolium.
FH	Key
FT	location/Qualifiers
FT	CDS 653..3560
FT	/*tag= a
FT	/product= CF-5 pathogen resistance gene
PN	MO9743429-A1.
PD	20-NOV-1997.
PF	08-MAY-1997. G01249.
PR	24-SEP-1996; GB-019924.
PR	09-MAY-1996; GB-009681.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI	Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR	WPI: 98-008895/01.
DR	P-PSDB: W41309.
PT	Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT	for production of transgenic plants resistant to pathogens e.g.
PT	tomato leaf mould C. fulvum in tomatoes
PS	Claim 3; Fig 1a; 75pp; English.
CC	This sequence is an example of the polynucleotide of the invention, and
CC	is able to confer pathogen resistance on a plant. It is one of two tomato
CC	CF-5 gene variants, which offer resistance against the pathogen
CC	Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC	the gene into plant cells and regenerating plants from the cells;
CC	asexually or sexually produced offspring can also be subsequently
CC	produced. Expression of the gene in plant cells can confer pathogen
CC	resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC	Oligonucleotides with sequences complementary to the gene or fragments
CC	of it, are useful in anti-sense techniques to reduce gene expression. The
CC	nucleic acids/polynucleotides are useful as hybridisation probes to
CC	identify other genes/fragments conferring pathogen resistance on plants
CC	e.g. Phytophthora resistance in potatoes. Homologues between Cf-5 and
CC	Cf-9 may be used to identify further resistance genes of this class.
90	Sequence 3979 BP: 1217 C; 717 C; 698 G; 1347 T;

Query Match 6.6%; Score 70.6; DB 1; Length 3979;  
Best Local Similarity 50.4%; Pred. No. 3.6e-10;  
Matches 172; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

OY	313	aacgcgcctactcgtgtggaatttgagggaattcaaacctctcggaaactcttgcgcctgag	372
Db	1944	AACACACTGTTTATGTATGTATCTTTACAAATAACACCTTTGGGCTCTTTCCTAATAA	2003
OY	373	cttgggaagcttgaacattacagatctagagctctacaaaaacacatcccaaggaact	432
Db	2004	ATAGGTTACCTGAGTCTCTTACTGAACTATTTTGGGTAAATACCTCTTAATGGCTT	2063
OY	433	ataccttcogaacttgaatactcgaagaactcatatgaattgagcttctgaacaaacat	492
Db	2064	ATTCTCGTTATTGGGGAATCTTAACAACTTCTCTAGGTTGATTTTACATATATAG	2123
OY	493	cttacaaggatagttcccaactctcttgggaataatgaagtcctcgtgcttlltacgctt	552
Db	2134	CTTCTCGCTCTAATTCCTGCTTCACTTTTGGCAATATAGAAATCTGCAAACTCTGTTTCTC	2183
OY	553	aatgaacaacgcataggaggggcacatccctcagaagcaactcactgcaatcccaagccttaa	612

Db 2184 AGAGTAAACATTCATTTGGGAAATTCCTCATTTTGTGTGCATTTGACATCACTGSA 2243

Oy 613 gtcttgatgctcgaagcaatgtttgtgtgacaaatccc 653

Db 2244 GTGTGTATATGTGCAGAAACATTTTGAAGGAAAAAGTTCC 2284

## RESULT 15

VI4519	standard; DNA; 3979 BP.
ID	VI4519
AC	20-MAY-1998 (first entry)
DI	CF-5 pathogen resistance gene variant #2.
DE	Tomato; CF-5 pathogen resistance gene; Cladosporium fulvum-5;
KW	tomato leaf mould; Phytophthora resistance; ss.
KV	Lycopersicon pimpinellifolium.
OS	Key location/Qualifiers
FE	653..3560
FT	CD5
FT	/tag= a
FT	/product= CF-5 pathogen resistance gene
PN	MO9743429-A1.
PD	20-NOV-1997.
PF	08-MAY-1997: G01249.
PR	24-SEP-1996: GB-019924.
PR	09-MAY-1996: GB-009681.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI	Dixon MS, Hatzixanthis K, Jones DA, Jones JDG;
DR	WPI: 98-008895/01.
DR	P-PSDB: W41310.
PT	Tomato gene Cf-5, confers resistance to Cladosporium fulvum - useful
PT	for production of transgenic plants resistant to pathogens e.g.
PT	tomato leaf mould C. fulvum in tomatoes
PS	Claim 6; Fig 1b; 75pp; English.
CC	This sequence is an example of the polynucleotide of the invention, and
CC	is able to confer pathogen resistance on a plant. It is one of two tomato
CC	Cf-5 gene variants, which offer resistance against the pathogen
CC	Cladosporium fulvum-5. Transgenic plants can be produced by incorporating
CC	the gene into plant cells and regenerating plants from the cells;
CC	asexually or sexually produced offspring can also be subsequently
CC	produced. Expression of the gene in plant cells can confer pathogen
CC	resistance on a plant e.g. to tomato leaf mould (C. fulvum) in tomatoes.
CC	Oligonucleotides with sequences complementary to the gene or fragments
CC	of it, are useful in anti-sense techniques to reduce gene expression. The
CC	nucleic acids/polynucleotides are useful as hybridisation probes to
CC	identify other genes/fragments conferring pathogen resistance on plants
CC	e.g. Phytophthora resistance in potatoes. Homologues between Cf-5 and
CC	Cf-9 may be used to identify further resistance genes of this class.
Sequence	3979 BP; 1218 A; 716 C; 658 G; 1347 T;

Query Match	6.68;	Score 70.6;	DB 1;	Length 3979;
Best Local Similarity	50.48;	Pred. No. 3.6e-10;		
Matches 172; Conservative	0;	Mismatches 169;	Indels 0;	Gaps 0;

QY	313	aacgcgcctactcgtcgtgatcttgggggaattcaaacctctcggacatcttgcgcctag	372
Db	1944	AACACCTGTTATGTATGTATCTTTACAAATAACACCTTTGGGCGATTTCTTGAA	2003
QY	373	cttgggaagccttgaacattcaacgtatctagagctctacaaaaaacacatccagaact	432
Db	2004	ATAGGTTACCTGAATTTCTTTACTGACACIATTTTGGGTAAATACCTCTTAATGGCTT	2065
QY	433	ataccttcgcgaacttggaaatctgaaagatctcatcagcttgcgatctgacacacaat	492
Db	2054	ATTCTGCTCATTTGGGGAACTAAACAATTGTTAGGTGATCTTAAATATACG	2122
QY	493	cttcaaggaatagctcccaactcttttgggaaattgaatctctcgtcttlltaagcctt	552
Db	2124	CTTTGTGGCTCTATTTCTCGTTCAATTTGGCAATATAGAAATCTGNAACCTGTGTTTTC	2183
QY	553	aatgaacaacccgatgaacggggccaatccctcagagcaactcaatccgaacttaaa	612
Db	2184	AGTAGTAAACGATCTCATTTGGGGAATTTCTCTTCAATTTGATGGAATTTTCACTACTGGA	2243



Qy 613 gtgttgatgtctcaagcaatgtttgtgtggaacaatccc 653  
Db 2244 GTGTTGTATATGTCGAGAAACAATTGGAAGGGAAGTTCC 2284

Search completed: June 23, 2000, 22:53:58  
Job time: 40634 sec